

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:10:18 ; Search time 9.63934 Seconds

(without alignments)
279,413 Million cell updates/sec

Title: US-09-673-166b-273
Perfect score: 140
Sequence: 1 GRQYKANSKFTIGTERGSLNVTATL 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 78:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	50.0	1315	1 BTCLTN	lentoxilysin (EC 3
2	52	37.1	500	2 S33979	gag polypeptide -
3	52	37.1	836	2 T18650	hypothetical prote
4	50	35.7	478	1 FOYWL	gag polypeptide -
5	50	35.7	497	1 FOLJND	gag polypeptide -
6	50	35.7	498	2 T09436	gag polypeptide -
7	50	35.7	500	1 FOYWL	gag polypeptide -
8	50	35.7	500	1 A44001	gag polypeptide -
9	50	35.7	500	1 FOYWL	gag polypeptide -
10	50	35.7	502	1 FOYWL	gag polypeptide -
11	50	35.7	512	1 FOYWL	gag polypeptide -
12	49	35.0	506	1 A38068	gag polypeptide -
13	49	35.0	506	1 A55236	gag polypeptide -
14	48.5	34.6	1060	2 S06286	gag polypeptide -
15	48.5	34.6	1086	2 S16752	gag polypeptide -
16	48.5	34.6	1701	2 A54498	gag polypeptide -
17	48.5	34.6	1701	2 A26868	gag polypeptide -
18	48.5	34.6	1726	1 SAZOGM	gag polypeptide -
19	48.5	34.6	1726	2 A45948	gag polypeptide -
20	48.5	34.6	3216	2 C90538	gag polypeptide -
21	48	34.3	83	2 B84174	gag polypeptide -
22	48	34.3	355	2 A97271	gag polypeptide -
23	48	34.3	975	2 A85318	gag polypeptide -
24	48	34.3	1126	2 T05762	gag polypeptide -
25	47.5	33.9	379	2 H90183	gag polypeptide -
26	47	33.6	505	2 T01667	gag polypeptide -
27	47	33.6	569	2 B82559	gag polypeptide -
28	47	33.6	674	2 T20571	gag polypeptide -
29	47	33.6	998	2 I41078	hemolysin - Escher

30	47	33.6	998	2 T00227	hemolysin A toxin
31	47	33.6	1161	2 G81186	conserved hypotet
32	47	33.6	1161	2 G81186	hypothetical prote
33	46	32.9	97	2 H64335	hypothetical prote
34	46	32.9	131	2 T12915	hypothetical prote
35	46	32.9	191	2 C64001	hypothetical prote
36	46	32.9	349	2 T43043	probable acetyl-Co
37	46	32.9	402	2 D81061	conserved hypotet
38	46	32.9	404	2 B81071	conserved hypotet
39	46	32.9	422	2 F81804	hypothetical prote
40	46	32.9	424	2 G97834	hemolysin homolog
41	46	32.9	447	2 H97146	hemolysin/Surfac
42	46	32.9	448	2 B64503	hypothetical prote
43	46	32.9	501	2 S54377	gag polypeptide -
44	46	32.9	823	2 G89007	protein F59B1.8 [I
45	46	32.9	999	2 S68689	glucose regulated

ALIGNMENTS

RESULT 1
BTCLTN
lentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N/Alternate names: tetanus neurotoxin
C/Species: Clostridium tetani
C/Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 03-Jun-2002
C/Accession: A25689; A25757; A25194; A60759; S69348; S09364
R/Eisel, U.; Jarasch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, M
EMBL J. 5, 2495-2502, 1986
A/Title: Tetanus toxin: primary structure, expression in E. coli, and homology with botu
A/Reference number: A25689; MUID:87053814; PMID:3536478
A/Accession: A25689
A/Molecule type: DNA
A/Residues: 1-1315 <EIS>
A/Cross-references: GB:X04436; NID:G40769; PIDN:CAA28033.1; PID:G40770
R/Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A/Title: The complete nucleotide sequence of tetanus toxin.
A/Reference number: A25757; MUID:87040747; PMID:3774547
A/Accession: A25757
A/Molecule type: DNA
A/Residues: 1-1315 <FAI>
A/Cross-references: GB:X06214; NID:G40773; PIDN:CAA2564.1; PID:G40774
R/Experimental source: strain CN3911
R/Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A/Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in E
A/Reference number: A25194; MUID:86085672; PMID:3510187
A/Accession: A25194
A/Molecule type: DNA
A/Residues: 743-1315 <FA2>
A/Cross-references: GB:M12739; NID:G144920; PIDN:AAA23282.1; PID:G144921
A/Accession: B25194
A/Molecule type: Protein
A/Residues: 865-894 <PA3>
R/Matsuda, M.; Iel, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A/Title: Isolation, purification, and characterization of fragment B, the NH-2-terminal
A/Reference number: A60759; MUID:90035436; PMID:2478476
A/Accession: A60759
A/Molecule type: Protein
A/Residues: 461-475 <WAT>
J. Demotz, S.; Iannavechia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 144, 394-402, 1989
A/Title: Determination of several DR-restricted tetanus toxin T cell epitopes.
A/Reference number: J50098; MUID:89093918; PMID:2463305
R/Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B.R
Nature 359, 832-835, 1992
A/Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic
A/Reference number: S27125; MUID:93063293; PMID:1331807
A/Contents: annotation

R/de Filipis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
Eur. J. Biochem. 229, 61-69, 1995
A/Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
A/Reference number: S69348; MUID:95262688; PMID:7744050
A/Accession: S69348
A/Molecule type: protein
A/Residues: 2-31 <DEF>
C/Comment: The source of this protein was an extrachromosomal plasmid.
C/Comment: The precursor is cleaved by endogenous proteinase activity to form light (fra
dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)
C/Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to ganglionic
C/Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized
presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy
C/Function:
A/Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in synapt
C/Superfamily: tetanus toxin
C/Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F/2-457/Product: tetroxylsyrin light chain (fragment A) #status predicted <TTH>
F/461-1315/Product: tetroxylsyrin heavy chain (fragment B,C) #status experimental <TTH>
F/461-864/Domain: channel forming (fragment B) #status predicted <TXB>
F/865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>
F/233-237/Binding site: zinc (His) #status predicted
F/234/Active site: Gln #status predicted

Query Match 50.0%; Score 70; DB 1; Length 1315;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 QYKANSKFIGITE 16
Db 830 QYKANSKFIGITE 843

RESULT 2
333979
gag polyprotein - human immunodeficiency virus type 1
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Aug-1999
C/Accession: S33979
R/Carlini, F.
Submitted to the EMBL Data Library, November 1991
A/Reference number: S33979
A/Accession: S33979
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-500 <CAR>
A/Cross-references: EMBL:Z11530; NID:960192; PIDN:CAA77621.1; PID:960193
C/Superfamily: AIDS-related virus gag polyprotein

Query Match 37.1%; Score 52; DB 2; Length 500;
Best Local Similarity 78.6%; Pred. No. 7.8;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 15 TERGRSLYNTVATL 28
Db 72 SEERSLYNTVATL 85

RESULT 3
T18650
Hypothetical protein B0035.12 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T18650
R/White, S.
Submitted to the EMBL Data Library, May 1996
A/Reference number: Z19002
A/Accession: T18650
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-836 <WIL>
A/Cross-references: EMBL:Z73102; PIDN:CAA97405.1; GSPDB:GN00022; CESP:B0035.12
A/Experimental source: clone B0035

C/Genetics:
A/Gene: CESP:B0035.12
A/Map position: 4
A/Introns: 28/3; 78/2; 134/3; 167/1; 299/3; 349/2; 550/1; 662/3; 700/3; 799/1

Query Match 37.1%; Score 52; DB 2; Length 836;
Best Local Similarity 52.2%; Pred. No. 13;
Matches 12; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

OY 6 KANSKFIGTERGRSLYNTVATL 28
Db 342 RARSNVINSAEGRSLYNTVATL 364

RESULT 4
FOVWVL
gag polyprotein - human immunodeficiency virus type 1 (isolate LV)
N/Alternate names: assemblin; core polyprotein; gag precursor
N/Contains: capsid antigen core protein p24CA; core protein p1; core protein p6; matrix ;
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 05-Dec-1998
C/Accession: A03948
R/Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.
Nature 313, 450-458, 1985
A/Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retroviru
A/Reference number: A93355; MUID:8511157; PMID:2982104
A/Accession: A03948
A/Molecule type: DNA
A/Residues: 1-478 <MOE>
C/Genetics:
A/Gene: gag
C/Superfamily: AIDS-related virus gag polyprotein
C/Keywords: AIDS; blocked amino end; core protein; immunodeficiency; lipoprotein; myrist
F/2-478/Product: gag precursor (assemblin) #status predicted <GAG>
F/2-132/Product: matrix antigen core protein p17MA #status predicted <P17>
F/20-32/Region: nuclear location signal
F/110-114/Region: nuclear location signal
F/133-363/Product: capsid antigen core protein p24CA #status predicted <P24>
F/364-377/Product: core protein p2 #status predicted <CP2>
F/378-452/Product: nucleocapsid core protein p7NC #status predicted <CP7>
F/392-405/Region: zinc finger CCHC motif
F/413-426/Region: zinc finger CCHC motif
F/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F/392,395,400,405/Binding site: zinc (Cys, Cys, His, Cys) #status experimental
F/413,416,421,426/Binding site: zinc (Cys, Cys, His, Cys) #status experimental

Query Match 35.7%; Score 50; DB 1; Length 478;
Best Local Similarity 78.6%; Pred. No. 15;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 15 TERGRSLYNTVATL 28
Db 72 SEERSLYNTVATL 85

RESULT 5
POLJND
gag polyprotein - human immunodeficiency virus type 1 (isolate NDK)
N/Alternate names: core polyprotein
N/Contains: core protein p15; core protein p17; core protein p24
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C/Accession: J00065
R/Spire, B.; Site, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe, A.;
Gene 81, 275-284, 1989
A/Title: Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the human immunc
A/Reference number: J00065; MUID:90034200; PMID:2806917
A/Accession: J00065
A/Molecule type: DNA
A/Residues: 1-497 <SP1>
A/Cross-references: GB:M27323; NID:9328154; PIDN:AAA44868.1; PID:9328157
C/Genetics:

433-448/Product: core protein p1 #status predicted <CP1>

Query Match	35.7%	Score 50	pp 1	total	acc

Query Match	35.7%	Score 50	pp 1	total	acc

Best Local Similarity 78.6%; Pred. No. 16;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 15 TERGRSLYNTVATL 28
: |||||
Db 72 SEELRSLYNTVATL 85

RESULT 10

FOVMA2
gag polyprotein - human immunodeficiency virus type 1 (isolate ARV-2)
N/Alternate names: core polyprotein
N/Contains: core protein p15; core protein p17; core protein p24
C/Species: human immunodeficiency virus type 1, HIV-1
A/Note: host Homo sapiens (man)
C/Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C/Accession: A03947
R/Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown-SH
Science 227, 484-492, 1985
A/Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).
A/Reference number: A04003; MUID:85090453; PMID:2578227
A/Accession: A03947
A/Residues: 1-502 <SAN>
A/Molecule type: DNA
C/Genetics: Cleavage sites that yield the mature core proteins remain to be determined.
A/Genes: gag
C/Superfamily: AIDS-related virus gag polyprotein
C/Keywords: AIDS; core protein; immunodeficiency; polyprotein
F:1-134/Product: core protein p17 #status predicted <P17>
F:135-193/Product: core protein p24 #status predicted <P24>
F:394-502/Product: core protein p15 #status predicted <P15>

Query Match 35.7%; Score 50; DB 1; Length 502;
Best Local Similarity 78.6%; Pred. No. 16;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
: |||||
Db 72 SEELRSLYNTVATL 85

RESULT 11

FOVWH3
gag polyprotein - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)
N/Alternate names: core polyprotein
N/Contains: core protein p15; core protein p17; core protein p24
C/Species: human immunodeficiency virus type 1, HIV-1
A/Note: host Homo sapiens (man)
C/Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C/Accession: A03945
R/Ratner, L.; Haseltine, W.; Patarca, R.; Liyak, K.J.; Starcich, B.; Josephs, S.F.; Dora
nderger, J.A.; Papas, T.S.; Ghayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.
Nature 313, 277-284, 1985
A/Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.
A/Reference number: A93353; MUID:85111123; PMID:2578615
A/Accession: A03945
A/Molecule type: DNA
A/Residues: 1-512 <RAT>
A/Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:9326383; PIDN:AAA442
A/Comment: Cleavage sites that yield the mature core proteins remain to be determined.
C/Genetics:
A/Genes: gag
C/Superfamily: AIDS-related virus gag polyprotein
C/Keywords: AIDS; core protein; immunodeficiency; polyprotein
F:1-132/Product: core protein p17 #status predicted <P17>
F:133-391/Product: core protein p24 #status predicted <P24>
F:392-512/Product: core protein p15 #status predicted <P15>

Query Match 35.7%; Score 50; DB 1; Length 512;
Best Local Similarity 78.6%; Pred. No. 16;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
: |||||
Db 72 SEELRSLYNTVATL 85

RESULT 12

A38068
gag polyprotein - human immunodeficiency virus type 1 (strain MN)
N/Alternate names: core polyprotein
N/Contains: core protein p1; core protein p17; core protein p2; core protein p24; core p
C/Species: human immunodeficiency virus type 1, HIV-1
A/Note: host Homo sapiens (man)
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-Feb-1999
C/Accession: A38068
R/Henderson, L.B.; Bowers, M.A.; Sowder II, R.C.; Serabyn, S.A.; Johnson, D.G.; Bees Jr.,
J. Virol. 66, 1856-1865, 1992
A/Title: Gag proteins of the highly replicative MN strain of human immunodeficiency virus
A/Reference number: A38068; MUID:92194415; PMID:1548743
A/Accession: A38068
A/Molecule type: protein
A/Residues: 1-506 <HEN>
C/Genetics:
A/Genes: gag
C/Superfamily: AIDS-related virus gag polyprotein
C/Keywords: AIDS; core protein; immunodeficiency; polyprotein
F:1-134/Product: core protein p17 #status experimental <P17>
F:135-365/Product: core protein p24 #status experimental <P24>
F:366-379/Product: core protein p2 #status experimental <P2>
F:380-434/Product: core protein p7 #status experimental <P7>
F:435-450/Product: core protein p1 #status experimental <P1>
F:451-506/Product: core protein p6 #status experimental <P6>

Query Match 35.0%; Score 49; DB 1; Length 506;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
: |||||
Db 71 SEERKSLYNTVATL 84

RESULT 13

A55236
kinesin-related protein KLP68D - fruit fly (Drosophila melanogaster)
N/Alternate names: kinesin-like protein 5; KLP5
C/Species: Drosophila melanogaster
C/Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 19-Jan-2001
C/Accession: A55236; E41298
J/Pesavento, P.A.; Stewart, R.J.; Goldstein, L.S.B.
J Cell Biol. 127, 1041-1048, 1994
A/Title: Characterization of the KLP68D kinesin-like protein in Drosophila: possible role
A/Reference number: A55236; MUID:95050960; PMID:7525600
A/Accession: A55236
A/Molecule type: mRNA
A/Residues: 1-784 <PEB>
A/Cross-references: GB:U15974; NID:9595912; PIDN:AAA69929.1; PID:9565090
R/Stewart, R.J.; Pesavento, P.A.; Woerpel, D.N.; Goldstein, L.S.B.
Proc. Natl. Acad. Sci. U.S.A. 88, 8470-8474, 1991
A/Title: Identification and partial characterization of six members of the kinesin super
A/Reference number: A41298; MUID:92020874; PMID:1924306
A/Accession: E41298
A/Molecule type: DNA
A/Residues: 'TC', '222-337', 'VRGV', <STE>
A/Cross-references: GB:M74431; NID:9157791; PIDN:AAA28658.1; PID:9157792
C/Genetics:
A/Genes: FlyBase:KLP68D; KLP5
A/Cross-references: FlyBase:FBgn0004381
C/Function:
A/Description: may be part of a motor protein that provides anterograde fast axonal tran
C/Superfamily: kinesin-related protein KIF3; kinesin motor domain homology
C/Keywords: ATP; coiled coil; microtubule binding; nucleotide binding; P-loop
F:1-349/Domain: head globular #status predicted <GGL>

F:20-350/Domain: kinesin motor domain homology <KMO>
 F:106-113/Region: nucleotide-binding motif A (P-loop)
 F:350-580/Domain: helical rod #status predicted <ROD>
 F:581-784/Domain: tail globular #status predicted <TEL>
 F:112/Binding site: ATP (lys) #status predicted

Query Match 35.0%; Score 49; DB 1; Length 784;
 Best Local Similarity 42.9%; Pred. No. 36;
 Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 QY 2 RQYIKANSKFI-GITERGRSLY 22
 Db 163 RDLKPKNSKHLVRRERGSVY 183

RESULT 14
 S06286
 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (st
 N/Alternate names: 190K protein; polymorphic schizont antigen
 C/Species: Plasmodium falciparum
 C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jun-2000
 C/Accession: S06286
 R/Certa, U.; Rotmann, D.; Matile, H.; Reber-Liske, R.
 EMBL J 6 4137-4142, 1987
 A/Title: A naturally occurring gene encoding the major surface antigen precursor p190 of
 A/Reference number: S06286; MUID:88166657; PMID:3327688
 A/Accession: S06286
 A/Status: not compared with conceptual translation
 A/Molecule type: DNA
 A/Residues: 1-1060 <CER>
 A/Superfamily: major merozoite surface antigen
 C/Keywords: surface antigen

Query Match 34.6%; Score 48.5; DB 2; Length 1060;
 Best Local Similarity 43.5%; Pred. No. 58;
 Matches 10; Conservative 7; Mismatches 5; Indels 1; Gaps 1;
 QY 2 RQYIKANSKFI-GITERGRSLY 23
 Db 982 KQFVKNSKVTITGLTQKNALN 1004

RESULT 15
 S16752
 major merozoite surface antigen MSA-1 - malaria parasite (Plasmodium falciparum) (fragme
 N/Alternate names: polymorphic schizont antigen p190
 C/Species: Plasmodium falciparum
 C/Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jun-2000
 C/Accession: A44865; S16752
 R/Olafsso, P.; Matile, H.; Certa, U.
 Exp. Parasitol. 74, 381-389, 1992
 A/Title: Plasmodium falciparum: the repetitive MSA-1 surface protein of the RO-71 isolat
 A/Reference number: A44865; MUID:92275047; PMID:1592091
 A/Accession: A44865
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1086 <OLA>
 A/Cross-references: EMBL:X61930
 A/Experimental source: isolate RO-71
 C/Genetics:
 A/Gene: MSA1
 C/Superfamily: major merozoite surface antigen
 C/Keywords: glycoprotein; merozoite; surface antigen; tandem repeat

Search completed: July 20, 2004, 06:34:08
 Job time : 11.6393 secs

Query Match 34.6%; Score 48.5; DB 2; Length 1086;
 Best Local Similarity 43.5%; Pred. No. 59;
 Matches 10; Conservative 7; Mismatches 5; Indels 1; Gaps 1;
 QY 2 RQYIKANSKFI-GITERGRSLY 23
 Db 1008 KQFVKNSKVTITGLTQKNALN 1030

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:17 ; Search time 5.5082 Seconds

(without alignments)
264.690 Million cell updates/sec

Title: US-09-673-166B-273
Perfect score: 140
Sequence: 1 GRQYKXNSKFIQTERGSLVNTVATL 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	70	50.0	1314	1	TEXT_CLOTE
2	54	38.6	499	1	GAG_HV1EL
3	51	36.4	499	1	GAG_HV1LM
4	51	36.4	513	1	MAD2_RHTLO
5	51	36.4	562	1	MM2S_HOMAN
6	50	35.7	388	1	GAG_HV1W2
7	50	35.7	496	1	GAG_HV1ND
8	50	35.7	499	1	GAG_HV1OY
9	50	35.7	499	1	GAG_HV1C4
10	50	35.7	499	1	GAG_HV1H2
11	50	35.7	499	1	GAG_HV1Y2
12	50	35.7	511	1	GAG_HV1A2
13	50	35.7	511	1	GAG_HV1B1
14	50	35.7	511	1	GAG_HV1B5
15	50	35.7	511	1	GAG_HV1BR
16	50	35.7	511	1	GAG_HV1PV
17	49	35.0	492	1	GAG_HV1U4
18	49	35.0	506	1	GAG_HV1MN
19	49	35.0	784	1	KL68_DROME
20	48.5	34.6	1682	1	MSP1_PLA3
21	48.5	34.6	1701	1	MSP1_PLA3
22	48.5	34.6	1701	1	MSP1_PLA3
23	48.5	34.6	1726	1	MSP1_PLA3
24	48.5	34.6	1726	1	MSP1_PLA3
25	47	33.6	504	1	GAG_HV1MA
26	47	33.6	532	1	TP6B_SULTO
27	46	32.9	97	1	TP6B_SULTO
28	46	32.9	191	1	Y096_HA2IN
29	46	32.9	448	1	Y096_HA2IN
30	46	32.9	500	1	Y096_HA2IN
31	46	32.9	500	1	Y096_HA2IN
32	46	32.9	999	1	OXRP_CRIBR
33	45	32.1	111	1	RU24_CHLNU

34	45	32.1	117	1	YCCL YEAST
35	45	32.1	146	1	DUT TREBA
36	45	32.1	151	1	Y228_AQJAE
37	45	32.1	152	1	GLB7_ARTSX
38	45	32.1	384	1	LE12_THRTN
39	45	32.1	530	1	TP6B_SULTO
40	45	32.1	530	1	TP6B_SULTO
41	44.5	31.8	1034	1	PMPE_CHLTR
42	44.5	31.8	1388	1	RPOB_XYUFA
43	44.5	31.8	1388	1	RPOB_XYUFA
44	44.5	31.8	2187	1	POLG_CXB20
45	44	31.4	127	1	PROF_SCHPO

ALIGNMENTS

RESULT 1
TEXT_CLOTE
ID TEXT_CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin) [Contains:
DE tetanus toxin light chain [tetanus toxin chain L]; Tetanus toxin heavy
DE chain [tetanus toxin chain H]].
GN TEXT OR CTPE0.
OS Clostridium tetani.
OG Plasmid pE88, and Plasmid 75 Kbp.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC PLASMID=75 Kbp;
RC MEDLINE=87053814; PubMed=3536478;
RA Eisel U., Jarasch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudek M., Habermann B., Niemann H.;
RT "Tetanus toxin: Primary structure, expression in E. coli, and
RT homology with botulinum toxins.";
RT BMO J. 5:2495-2502(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CN911; PLASMID=75 Kbp;
RC MEDLINE=87040747; PubMed=3774547;
RA Falzweather N.F., Lyness V.A.;
RT "The complete nucleotide sequence of tetanus toxin.";
RT Nucleic Acids Res. 14:7809-7812(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88; PLASMID=PE88;
RC MEDLINE=22457253; PubMed=12552129;
RA Bruggemann H., Baumer S., Fricke W.F., Wietzer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
RN [4]
RP SEQUENCE OF 742-1314 FROM N.A.
RC PLASMID=75 Kbp;
RC MEDLINE=86085672; PubMed=3510187;
RA Falzweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT fragment C in Escherichia coli.";
RT J. Bacteriol. 165:21-27(1986).
RN [5]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RC MEDLINE=90201034; PubMed=2108021;
RA Krieglstein K., Henschen A., Weller U., Habermann B.;
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
RT in tetanus toxin.";

RL Eur. J. Biochem. 188:39-45(1990).
RN [6]
RP PARTIAL SEQUENCE.
RX MEDLINE=92037649; PubMed=1935979;
RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
RT "Limited proteolysis of tetanus toxin. Relation to activity and
RT identification of cleavage sites."
RL Eur. J. Biochem. 202:41-51(1991).
RN [7]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93010948; PubMed=1396558;
RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
RA Montecucco C.;
RT "Tetanus toxin is a zinc protein and its inhibition of
RT neurotransmitter release and protease activity depend on zinc."
RL EMBJ. 11:3577-3583(1992).
RN [8]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=93063293; PubMed=131807;
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA Dasgupta B.R., Montecucco C.;
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin."
RL Nature 359:832-835(1992).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
RX MEDLINE=97475217; PubMed=9334741;
RA Unland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
RA Sax M.;
RT "Structure of the receptor binding fragment HC of tetanus
RT neurotoxin."
RL Nat. Struct. Biol. 4:788-792(1997).
CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
CC BOND OF SYNAPTOSOMAL-2.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-|-Phe-77 bond in
CC synaptobrevin 2.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (by similarity).
CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
CC AND ARE NON-TOXIC AFTER SEPARATION.
CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
CC GLYCOSYLASE RECEPTORS.
CC -1- SIMILARITY: Belongs to peptidase family M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X0436; CAA28033.1; -;
DR EMBL; X06214; CAA29564.1; -;
DR EMBL; AF528097; AA037454.1; -;
CC EMBL; M12739; AAA23282.1; -;
DR PIR; A25689; BTCLTN.
DR PDB; 1AF9; 29-APR-98.
DR PDB; 1ABD; 14-OCT-98.
DR PDB; 1DOH; 27-MAR-00.
DR PDB; 1DPO; 24-MAR-00.
DR PDB; 1DIW; 24-MAR-00.
DR PDB; 1DLI; 24-MAR-00.
DR PDB; 1FV3; 05-SEP-01.
DR MEROPS; M27.001; -;
DR InterPro; IPR008985; Cons like Jec_g1.
DR InterPro; IPR002160; Kunitz legume.
DR InterPro; IPR006025; Pept_M_Zn_BS.

DR InterPro; IPR000395; Peptidase_M27.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOKILYSIN.
DR ProDom; PD001963; Bontokilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolyase; Metalloprotease; Zinc; Plasmid;
KW 3D-structure; Complete proteome.
FT INIT MET 0
FT CHAIN 1 456
FT METAL 1314
FT ACT SITE 232 232
FT ACT SITE 233 233
FT METAL 236 236
FT TRANSMEM 226 246
FT TRANSMEM 669 689
FT DISULFID 438 466
FT DISULFID 1076 1092
FT HELIX 876 882
FT TURN 883 883
FT TURN 884 891
FT TURN 892 893
FT TURN 894 897
FT STRAND 904 907
FT TURN 909 910
FT STRAND 912 915
FT STRAND 920 925
FT TURN 928 929
FT STRAND 932 935
FT STRAND 938 940
FT TURN 941 946
FT STRAND 949 956
FT HELIX 962 968
FT TURN 969 970
FT STRAND 972 977
FT STRAND 980 981
FT STRAND 983 985
FT HELIX 987 995
FT STRAND 996 997
FT TURN 998 1004
FT STRAND 1006 1007
FT TURN 1010 1016
FT STRAND 1020 1020
FT TURN 1021 1022
FT STRAND 1031 1037
FT TURN 1039 1040
FT STRAND 1042 1047
FT TURN 1048 1049
FT STRAND 1050 1056
FT TURN 1058 1059
FT STRAND 1068 1074
FT TURN 1079 1080
FT STRAND 1082 1091
FT HELIX 1097 1105
FT TURN 1106 1107
FT STRAND 1112 1112
FT STRAND 1114 1114
FT TURN 1116 1117
FT STRAND 1120 1120
FT STRAND 1122 1122
FT TURN 1123 1124
FT STRAND 1127 1131
FT HELIX 1132 1134
FT TURN 1135 1136
FT STRAND 1141 1141
FT TURN 1144 1145
FT STRAND 1148 1152
FT STRAND 1155 1158
FT TURN 1159 1162
FT STRAND 1163 1166
FT STRAND 1173 1178
FT TURN 1184 1185
FT STRAND 1188 1188
FT STRAND 1190 1190
TETANUS TOXIN LIGHT CHAIN.
ZINC (CATALYTIC) (BY SIMILARITY).
BY SIMILARITY.
ZINC (CATALYTIC) (BY SIMILARITY).
POTENTIAL.
INTERCHAIN.

Query Match 50.0%; Score 70; DB 1; Length 1314;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QYKANSKFIGITE 16
 |||||

DB 829 QYKANSKFIGITE 842

RESULT 2

GAG_HV1EL
 ID GAG_HV1EL STANDARD; PRT; 499 AA.
 AC P04592;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GAG polypeptide [Contains: Core proteins p17, p24, p2, p7, p1, p6].
 GN GAG.
 OS Human immunodeficiency virus type 1 (ELI isolate) (HTV-1).
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86245056; PubMed=2424612;
 RA Allison M., Main-Hobson S., Montagnier L., Sonigo P.;
 RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
 of two isolates from African patients";
 RL Cell 46:63-74 (1986).
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
 CC -1- PTM: The p24 protein is phosphorylated.
 CC -1- SIMILARITY: Contains 2 CCHC-type zinc fingers.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; K03454; AAA44324.1; -;
 DR EMBL; A07108; CA000611.1; -;
 DR HSSP; P05888; 1AAR.
 DR HTV; K03454; GAGSELI.
 DR InterPro; IPR000721; Gag_p24.
 DR InterPro; IPR008916; Retrov_capsid_C.
 DR InterPro; IPR008919; Retrov_capsid_N.
 DR InterPro; IPR000071; Retrovir_p17.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00540; Gag_p17; 1.
 DR Pfam; PF00607; Gag_p24; 1.
 DR Pfam; PF00098; Zf_CCHC; 2.
 DR PRINTS; PR00939; C2HCZNFINGER.
 DR PRINTS; PR00243; HIVMATRIX.
 DR SMART; SM00343; Znf_C2HC; 2.
 DR SMART; PS50158; Zf_CCHC; 2.
 DR PROSITE; PS50158; Zf_CCHC; 2.
 DR AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
 Zinc-finger; Repeat; Lipoprotein.
 KM Zinc-finger; Repeat; Lipoprotein.
 FT INIT MET 0
 FT CHAIN 1 131
 FT CHAIN 132 362
 FT CHAIN 363 377
 FT CHAIN 378 431
 FT CHAIN 432 447
 FT CHAIN 448 499
 FT CHAIN 499 500
 FT ZN_FING 390 407
 FT CCHC-TYPE 1.

FT ZN_FING 411 428 CCHC-TYPE 2.
 FT LIPID 1 1 N-myristoyl glycine (in host) (By
 FT SEQUENCE 499 AA; 55582 MW; 8A1785A59EARD08D CRC64;
 Query Match 38.6%; Score 54; DB 1; Length 499;
 Best Local Similarity 85.7%; Pred. No. 1.4;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 |||||

DB 71 TERGRSLYNTVATL 84

RESULT 3

GAG_HV1LW
 ID GAG_HV1LW STANDARD; PRT; 499 AA.
 AC Q70622;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GAG polypeptide [Contains: Core proteins p17, p24, p2, p7, p1, p6].
 GN GAG.
 OS Human immunodeficiency virus type 1 (LM12.3 isolate) (HTV-1).
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=82834;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95127297; PubMed=7826699;
 RA Reitz M.S., Jr., Hall J., Robert-Guroff M., Lautenberger J., Hahn B.M.,
 RA Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
 RT "Viral variability and serum antibody response in a laboratory worker
 infected with HIV type 1 (HTLV type IIb).";
 RL AIDS Res. Hum. Retroviruses 10:1143-1155 (1994).
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
 CC -1- PTM: The p24 protein is phosphorylated.
 CC -1- SIMILARITY: Contains 2 CCHC-type zinc fingers.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U12055; AAA76686.1; -;
 DR PDB; 1HVN; 31-JAN-94.
 DR PDB; 1HVO; 31-JAN-94.
 DR PDB; 2ZNF; 15-JUL-92.
 DR InterPro; IPR000721; Gag_p24.
 DR InterPro; IPR008916; Retrov_capsid_C.
 DR InterPro; IPR008919; Retrov_capsid_N.
 DR InterPro; IPR000071; Retrovir_p17.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00540; Gag_p17; 1.
 DR Pfam; PF00607; Gag_p24; 1.
 DR Pfam; PF00098; Zf_CCHC; 2.
 DR PRINTS; PR00939; C2HCZNFINGER.
 DR PRINTS; PR00243; HIVMATRIX.
 DR SMART; SM00343; Znf_C2HC; 2.
 DR SMART; PS50158; Zf_CCHC; 2.
 DR AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
 Zinc-finger; Repeat; 3D-structure; Lipoprotein.
 KM Zinc-finger; Repeat; 3D-structure; Lipoprotein.
 FT INIT MET 0
 FT CHAIN 1 131
 FT CHAIN 132 362
 FT CHAIN 363 377
 FT CHAIN 378 431
 FT CHAIN 432 447
 FT CHAIN 448 499
 FT CHAIN 499 500
 FT ZN_FING 390 407
 FT CCHC-TYPE 1.

```

FT CHAIN 363 376 CORE PROTEIN P2.
FT CHAIN 377 431 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 432 447 CORE PROTEIN P1.
FT CHAIN 448 499 CORE PROTEIN P6.
FT ZN_FING 389 406 CCHC-TYPE 1.
FT ZN_FING 410 427 CCHC-TYPE 2.
FT LIPID 1 1 N-myristoyl glycine (in host) (By
FT SEQUENCE 499 AA; 55772 MW; AA2F683546DC0A9 CRC64;
  similarity).
Query Match 36.4%; Score 51; DB 1; Length 499;
Best Local Similarity 78.6%; Pred. NO. 4.1;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
Db 71 SEGRSLYNTVATL 84

RESULT 4
NADB_RHILO STANDARD; PRT; 513 AA.
AC 0984V8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE L-aspartate oxidase (EC 1.4.3.16) (LASPO) (Quinolinate synthetase B).
NADB OR ML5834.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kato T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idehara K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimoto S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RA "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti";
RL DNA Res. 7:331-338(2000).
CC -1- FUNCTION: Catalyzes the oxidation of L-aspartate to
CC L-aspartate.
CC -1- CATALYTIC ACTIVITY: L-aspartate + H(2)O + O(2) = oxaloacetate +
CC NH(3) + H(2)O(2).
CC -1- COFACTOR: FAD.
CC -1- PATHWAY: NAD biosynthesis; aspartate to NAMN; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE FAD-DEPENDENT OXIDOREDUCTASE FAMILY 2.
CC NADB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL contribution
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
CC EMBL, AP003007; BAB52214.1; -
CC InterPro; IPR003953; FAD_bind2.
CC InterPro; IPR005288; NADB.
CC InterPro; IPR004112; Succ_DH flav_C.
CC Pfam; PF00890; FAD_binding_2; 1.
CC Pfam; PF02910; succ_DH flav_C; 1.
CC TIGRPFAM; TIGR00551; nadb; 1.
CC Pyridine nucleotide biosynthesis; Oxidoreductase; Flavoprotein; FAD;
CC Complete proteome.
CC NP_BIND 13 27 FAD (AMP PART) (POTENTIAL).
CC ACT_SITE 233 233 BY SIMILARITY.

```

```

FT ACT SITE 252 252 BY SIMILARITY.
SQ SEQUENCE 513 AA; 53469 MW; 6DA2B4B63F4200BC CRC64;
Query Match 36.4%; Score 51; DB 1; Length 513;
Best Local Similarity 55.0%; Pred. NO. 4.2;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 9 SKRTGTERGRSLYNTVATL 28
Db 422 SNALGIVRNKSLCDIVATL 441

RESULT 5
NM25_HUMAN STANDARD; PRT; 562 AA.
ID NM25_HUMAN
AC Q9NP42; Q9H3Q0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Matrix metalloproteinase-25 precursor (EC 3.4.24.-) (MMP-25)
DE (Membrane-type matrix metalloproteinase 6) (MT6-MMP 6) (Membrane-type-6
DE matrix metalloproteinase) (MT6-MMP) (Leukolysin).
GN MMP25 OR MT6MMP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal liver;
RX MEDLINE=2016815; PubMed=10706098;
RA Velasco G., Cal S., Merlos-Suarez A., Ferrando A.A., Alvarez S.,
RA Nakano A., Arribas J., Lopez-Otin C.;
RA "Human MT6-matrix metalloproteinase: identification, progelatinase A
RT activation, and expression in brain tumors.";
RL Cancer Res. 60:877-882(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20092460; PubMed=10628838;
RA Pei D.O.;
RT "Leukolysin/MMP25/MT6-MMP: a novel matrix metalloproteinase
RT specifically expressed in the leukocyte lineage.";
RL Cell Res. 9:291-303(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Kojima S., Itoh Y., Matsumoto S., Seiki M.;
RT "Membrane-type 6 matrix metalloproteinase (MT6-MMP, MMP-25) is the
RT second glycosyl-phosphatidyl inositol (GPI)-anchored MMP.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA De Santis-Vis B.M., Clair-Moninot V.A., Lambert C.A., Vandervliet B.,
RA Pin J.J., Chalus L., Alt-Yahia S., Caux C., Richelle-Nusgens B.V.,
RA Fossiez F., Lebecque S.;
RT "Molecular cloning of a novel human membrane-type matrix
RT metalloproteinase (MT-MMP) predominantly expressed in dendritic
RT cells.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP GPI-ANCHOR.
RX MEDLINE=20487129; PubMed=11034316;
RA Kojima S.-I., Itoh Y., Matsumoto S.-I., Masuno Y., Seiki M.;
RT "Membrane-type 6 matrix metalloproteinase (MT6-MMP, MMP-25) is the
RT second glycosyl-phosphatidyl inositol (GPI)-anchored MMP.";
RL FEBS Lett. 480:142-146(2000).
CC -1- FUNCTION: May activate progelatinase A.
CC -1- COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI anchor.
CC -1- TISSUE SPECIFICITY: Expressed predominantly in leukocytes, lung
CC and spleen. Expressed also in colon carcinoma, astrocytoma and
CC glioblastomas.
CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).

```

```

CC -1- SIMILARITY: Belongs to peptidase family M10A.
CC -1- SIMILARITY: Contains 1 hemopexin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ239053; CAB94713.1; -
DR EMBL; AF145442; AAF66697.2; -
DR EMBL; AF185270; AAG17007.1; -
DR EMBL; AB042328; BAB20584.1; -
DR EMBL; AJ272137; CAC03490.1; -
DR HSSP; Q02853; 1HV5.
DR MEROPS; M10.024; -.
DR Genew; HGNC:14246; MMP25.
DR GO; GO:0005578; C:extracellular matrix; NAS.
DR GO; GO:0016020; C:membrane; NAS.
DR GO; GO:0006994; P:inflammatory response; NAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Peptidase_M.
DR Pfam; PF00043; hemopexin; 4.
DR Pfam; PF03933; Peptidase_M10_N; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZINC; 1.
DR PROSITE; PS00024; HEMOPEXIN; FALSE NEG.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydroxylase; Metalloprotease; Zinc; Calcium; Signal; Zymogen;
KW GFI-anchor; Extracellular matrix; Lipoprotein.
FT SIGNAL; 1 21
FT PROPEP; 22 107 BY SIMILARITY.
FT CHAIN; 108 539 MATRIX METALLOPROTEINASE-25.
FT PROPEP; 540 562 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN; 314 511 HEMOPEXIN-LIKE.
FT SITE; 90 90 CYSTEINE SWITCH (POTENTIAL).
FT METAL; 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT SITE; 234 234 BY SIMILARITY.
FT METAL; 237 237 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL; 243 243 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID; 317 508 BY SIMILARITY.
FT LIPID; 539 539 GPI-anchor amidated alanine (Potential).
FT DOMAIN; 103 107 POLY-ARG.
FT CHAIN; 549 555 POLY-LEU.
FT CONFLICT; 47 47 P -> R (IN REF. 3).
SQ SEQUENCE 562 AA; 62554 MW; A6A50AE05D969C64 CRC64;
Query Match 36.4%; Score 51; DB 1; Length 562;
Best Local Similarity 44.4%; Pred. No. 4.7;
Matches 12; Conservative 3; Mismatches 12; Indels 0; Gaps 0;
QY 2 RQYIKANSKFGITGRSLYNTVATL 28
Db 59 RDAIKVMORFAGLPETGRMDPGTVATM 85

```

```

RESULT 6
GAG_HY1W2 STANDARD; PRT; 388 AA.
AC P05889;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polypeptide [Contains: Core proteins P17, P24, P2, P7, P1, P6]
GN (Fragment).

```

```

OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCHI_taxid=11705;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=86235450; PubMed=3012778;
CC Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
CC Salahuddin S.Z., Wong-Staal P., Gallo R.C., Parks E.S., Parks W.P.;
CC "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
CC at risk for AIDS."
CC Science 232:1548-1553(1986).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC -1- BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
CC WAS PERINATALLY INFECTED BY HER MOTHER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03457; AAB12988.1; -
DR PDB; 1FGJ; 01-APR-97.
DR HIV; K03457; GAG$WMJ2.
DR InterPro; IPR000721; Gag_p24.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR000071; Retrovir_p17.
DR Pfam; PF00540; Gag_p17; 1.
DR PRINTS; PR00234; HIVMATRIX.
DR AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; 3D-structure; Lipoprotein.
FT INIT MET; 0 0 BY SIMILARITY.
FT CHAIN; 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN; 132 362 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN; 363 376 CORE PROTEIN P2.
FT CHAIN; 377 >388 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT LIPID; 1 1 N-myristoyl glycine (in host) (By
FT similarity).
SQ SEQUENCE 388 AA; 43314 MW; EF885BE10BCF7804 CRC64;
Query Match 35.7%; Score 50; DB 1; Length 388;
Best Local Similarity 78.6%; Pred. No. 4.5;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 15 TERGSLYNTVATL 28
Db 71 SEIRSLYNTVATL 84

```

```

RESULT 7
GAG_HY1ND STANDARD; PRT; 496 AA.
AC P18600;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polypeptide [Contains: Core proteins P17, P24, P2, P7, P1, P6].
GN Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

```

NCBI_TaxID=11695;
[1]
SEQUENCE FROM N.A.
MEDLINE=90034200; PubMed=2806917;
RA Spire B., Spire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,
Hampe A., Chermann J.C.,
RT "Nucleotide sequence of HIV-1 NDK: a highly cytopathic strain of the
human immunodeficiency virus.";
RL Gene 81.275-284(1989).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- PTM: The p24 protein is phosphorylated.
CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
CC -1- SIMILARITY: Contains 2 CCHC-type zinc fingers.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M27323; AAA4468.1; -.
CC PIR; J00065; FOLND.
CC PDB; 1EQ1; 30-SEP-94.
CC HIV; M27323; GAGSNDK.
CC InterPro; IPR000721; Gag_P24.
CC InterPro; IPR008916; Retrov_capsid_C.
CC InterPro; IPR008919; Retrov_P17.
CC InterPro; IPR000771; Retrov_P17.
CC InterPro; IPR001878; Znf_CCHC.
CC Pfam; PF00540; Gag_P17; 1.
CC Pfam; PF00607; Gag_P24; 1.
CC Pfam; PF00098; Zf_CCHC; 2.
CC PRINTS; PR00939; C2HCZNFINGER.
CC PRINTS; PR00234; HIVMATRIX.
CC SMART; SM00343; Znf_CCHC; 2.
CC PROSITE; PS50158; Zf_CCHC; 2.
CC AIDS; Core protein; Polypeptide; Myristate; Phosphorylation;
CC Zinc-finger; Repeat; 3D-structure; Lipoprotein.
CC INIT MET 0
CC CHAIN 1 128 CORE PROTEIN P17 (MATRIX PROTEIN).
CC CHAIN 129 359 CORE PROTEIN P24 (CORE ANTIGEN).
CC CHAIN 360 374 CORE PROTEIN P2. (NUCLEOCAPSID PROTEIN).
CC CHAIN 375 429 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
CC CHAIN 430 445 CORE PROTEIN P1.
CC CHAIN 446 496 CORE PROTEIN P6.
CC CHAIN 387 404 CCHC-TYPE 1.
CC ZN_FING 387 404 CCHC-TYPE 1.
CC ZN_FING 408 425 CCHC-TYPE 2.
CC LIPID 1 N-myristoyl glycine (in host) (By
CC similarity).
CC
CC SEQUENCE 496 AA; 55152 MW; F50A42B42CCF20AA CRC64;
Query Match 35.7%; Score 50; DB 1; Length 496;
Best Local Similarity 78.6%; Pred. No. 5.8;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

01-FEB-1991 (Rel. 17, Created)
01-FEB-1994 (Rel. 28, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].
GN GAG.
OS Human immunodeficiency virus type 1 (OVI isolate) (HIV-1).
OC Viruses; Retrov. Viruses; Retroviridae; Lentivirus.
OK NCBI_TaxID=11699;
RN
RN SEQUENCE FROM N.A.
RP MEDLINE=90148544; PubMed=2559749;
RA Huet T., Daza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
individual presenting an atypical western blot.";
RL AIDS 3:707-715(1989).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- PTM: The p24 protein is phosphorylated.
CC -1- MISCELLANEOUS: THE OVI ISOLATE WAS TAKEN FROM THE BLOOD OF A
HEALTHY GABONESE INDIVIDUAL.
CC -1- SIMILARITY: Contains 2 CCHC-type zinc fingers.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M26727; AA83391.1; -.
CC HSSP; P05888; IAAF.
CC HIV; M26727; GAGSOYI.
CC InterPro; IPR000721; Gag_P24.
CC InterPro; IPR008916; Retrov_capsid_C.
CC InterPro; IPR008919; Retrov_capsid_N.
CC InterPro; IPR000771; Retrov_P17.
CC InterPro; IPR001878; Znf_CCHC.
CC Pfam; PF00540; Gag_P17; 1.
CC Pfam; PF00607; Gag_P24; 1.
CC Pfam; PF00098; Zf_CCHC; 2.
CC PRINTS; PR00939; C2HCZNFINGER.
CC PRINTS; PR00234; HIVMATRIX.
CC SMART; SM00343; Znf_CCHC; 2.
CC PROSITE; PS50158; Zf_CCHC; 2.
CC AIDS; Core protein; Polypeptide; Myristate; Phosphorylation;
CC Zinc-finger; Repeat; Lipoprotein.
CC INIT MET 0
CC CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).
CC CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).
CC CHAIN 363 376 CORE PROTEIN P2. (NUCLEOCAPSID PROTEIN).
CC CHAIN 377 430 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
CC CHAIN 431 446 CORE PROTEIN P1.
CC CHAIN 447 498 CORE PROTEIN P6.
CC CHAIN 388 405 CCHC-TYPE 1.
CC ZN_FING 388 405 CCHC-TYPE 1.
CC ZN_FING 409 426 CCHC-TYPE 2.
CC LIPID 1 N-myristoyl glycine (in host) (By
CC similarity).
CC
CC SEQUENCE 498 AA; 55451 MW; BDFE20658DEB20B1 CRC64;

Query Match 35.7%; Score 50; DB 1; Length 498;
Best Local Similarity 78.6%; Pred. No. 5.9;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 15 TERGSLVNTVATL 28
DB 71 SEIRSLVNTVATL 84


```

RESULT 9
GAG_HV1C4
ID GAG_HV1C4 STANDARD; PRT; 499 AA.
AC P05887;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polypeptide [Contains: Core proteins P17, P24, P2, P7, P1, P6].
GN GAG.
OS Human immunodeficiency virus type 1 (CDC-451 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11687;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87041461; PubMed=3490666;
RA Desai S.M., Kalyanaram V.S., Casey J.M., Srinivasan A.,
RA Andersen P.R., Devare S.G.;
RT "Molecular cloning and primary nucleotide sequence analysis of a
RT distinct human immunodeficiency virus isolate reveal significant
RT divergence in its genomic sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 83:8380-8384 (1986).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- PTM: The p24 protein is phosphorylated.
CC -1- SIMILARITY: Contains 2 CCHC-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13136; AAA44306.1; -.
DR PIR; A25523; F0VWH4.
DR HSSP; P05888; 1A6F.
DR HIV; M13136; GAGSCDC45.
DR InterPro; IPR0000721; Gag_p24.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR000071; Retrovir_p17.
DR Pfam; PF00540; Gag_p17; 1.
DR Pfam; PF00607; Gag_p24; 1.
DR Pfam; PF00607; Zf_CCHC; 2.
DR PRINTS; PR00098; ZF-CCHC; 2.
DR PRINTS; PR00939; C2HC2NFINGER.
DR PRINTS; PR00234; HIVMATRIX.
DR SMART; SM00343; Znf_C2HC; 2.
DR PROSITE; PS50158; ZF_CCHC; 2.
KW Aids; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat; Lipoprotein.
FT INIT MET 0
FT CHAIN 1 131 BY SIMILARITY.
FT CHAIN 132 362 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 363 376 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 377 431 CORE PROTEIN P2.
FT CHAIN 432 447 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 448 499 CORE PROTEIN P1.
FT ZN_FING 389 406 CORE PROTEIN P6.
FT ZN_FING 410 427 CCHC-TYPE 1.
FT LIPID 1 1 N-myristoyl glycine (in host) (by
SQ SEQUENCE 499 AA; 55796 MW; 023CA76C9C6F22AD CRC64;
Query Match 35.7%; Score 50; DB 1; Length 499;
Best Local Similarity 78.6%; Pred. No. 5.9;

```

```

Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Oy 15 TERGRSLYNTVATL 28
Db 71 SEELRSLYNTVATL 84
RESULT 10
GAG_HV1H2
ID GAG_HV1H2 STANDARD; PRT; 499 AA.
AC P04591;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polypeptide [Contains: Core proteins P17, P24, P2, P7, P1, P6].
GN GAG.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87292916; PubMed=3040055;
RA Ratner L., Fisher A., Zagodszinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequences of functional clones of the AIDS
RT virus."
RL Aids Res. Hum. Retroviruses 3:57-69 (1987).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- PTM: The p24 protein is phosphorylated.
CC -1- SIMILARITY: Contains 2 CCHC-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03455; AAB50258.1; -.
DR PDB; 1BK1; 11-DEC-00.
DR PDB; 1ESK; 21-APR-00.
DR PDB; 1NCP; 31-OCT-93.
DR PDB; 1TAM; 12-NOV-96.
DR HIV; K03455; GAGSHXB2.
DR InterPro; IPR000721; Gag_p24.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR000071; Retrovir_p17.
DR Pfam; PF00540; Gag_p17; 1.
DR Pfam; PF00607; Gag_p24; 1.
DR Pfam; PF00607; Zf_CCHC; 2.
DR PRINTS; PR00098; ZF-CCHC; 2.
DR PRINTS; PR00939; C2HC2NFINGER.
DR PRINTS; PR00234; HIVMATRIX.
DR SMART; SM00343; Znf_C2HC; 2.
DR PROSITE; PS50158; ZF_CCHC; 2.
KW Aids; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat; 3D-structure; Lipoprotein.
FT INIT MET 0
FT CHAIN 1 131 BY SIMILARITY.
FT CHAIN 132 362 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 363 376 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 377 431 CORE PROTEIN P2.
FT CHAIN 432 447 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 448 499 CORE PROTEIN P1.
FT ZN_FING 389 406 CORE PROTEIN P6.
FT CCHC-TYPE 1.

```

FT ZN_FING 410 427 CCHC-TYPE 2.
 FT LIPID 1 1 N-myristoyl glycine (in host) (By
 FT SEQUENCE 499 AA; 55798 MW; 774C384D6EACB108 CRC64;
 Query Match 35.7%; Score 50; DB 1; Length 499;
 Best Local Similarity 78.6%; Pred. No. 5.9;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGSLYNTVATL 28
 Db 71 SEELRLSYNTVATL 84

RESULT 11
 GAG_HV1A2 STANDARD; PRT; 499 AA.
 AC P35962;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GAG polyprotein [Contains: Core proteins P17, P24, P2, P7, P1, P6].
 GN GAG
 OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=36377;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93021387; PubMed=1404605;
 RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
 Shaw G.M.;
 RT "Complete nucleotide sequence, genome organization, and biological
 RT properties of human immunodeficiency virus type 1 in vivo: evidence
 RT for limited defectiveness and complementation.";
 RL J. Virol. 66:6587-6600 (1992).
 CC -!- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
 CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
 CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
 CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
 CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
 CC -!- P17: The p24 protein is phosphorylated.
 CC -!- SIMILARITY: Contains 2 CCHC-type zinc fingers.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M93258; -!- NOT_ANNOTATED_CDS.
 CC PIR: A44001; A44001.
 CC PDB: 1F6U; 09-OCT-00.
 CC InterPro: IPR000721; Gag_p24.
 CC InterPro: IPR008916; Retrov_capsid_C.
 CC InterPro: IPR008919; Retrov_capsid_N.
 CC InterPro: IPR000071; Retrovir_p17.
 CC InterPro: IPR001878; Znf_CCHC.
 CC Pfam: PF00540; Gag_p17; 1.
 CC Pfam: PF00607; Gag_p24; 1.
 CC Pfam: PF00098; zf-CCHC; 2.
 CC PRINTS: PR00939; C2HCZNFINGER.
 CC PRINTS: PR00234; HIVMATRIX.
 CC SMART: SM00343; ZNF_C2HC; 2.
 CC PROSITE: PS50158; ZF_CCHC; 2.
 CC AIDs; Core protein; Polyprotein; Myristate; Phosphorylation;
 CC zinc-finger; Repeat; 3D-structure; Lipoprotein.
 CC INIT MET 0
 CC CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).
 FT

FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).
 FT CHAIN 363 376 CORE PROTEIN P2. (NUCLEOCAPSID PROTEIN).
 FT CHAIN 377 431 CORE PROTEIN P7.
 FT CHAIN 432 447 CORE PROTEIN P1.
 FT CHAIN 448 499 CORE PROTEIN P6.
 FT ZN_FING 389 406 CCHC-TYPE 1.
 FT ZN_FING 410 427 CCHC-TYPE 2.
 FT LIPID 1 1 N-myristoyl glycine (in host) (By
 FT SEQUENCE 499 AA; 55660 MW; 278B65F5405CD99 CRC64;
 Query Match 35.7%; Score 50; DB 1; Length 499;
 Best Local Similarity 78.6%; Pred. No. 5.9;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGSLYNTVATL 28
 Db 71 SEELRLSYNTVATL 84

RESULT 12
 GAG_HV1A2 STANDARD; PRT; 501 AA.
 AC P03349;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GAG polyprotein [Contains: Core proteins P17, P24, P2, P7, P1, P6].
 GN GAG.
 OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85090453; PubMed=2578227;
 RA Sanchez-Pescador R., Power M.D., Barr P.J., Steiner K.S.,
 Stempelen M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
 Levy J.A., Dina D., Luciw P.A.;
 RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
 RT (ARV-2)."
 RL Science 227:484-492 (1985).
 CC -!- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
 CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
 CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
 CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
 CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
 CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
 CC -!- P17: The p24 protein is phosphorylated.
 CC -!- SIMILARITY: Contains 2 CCHC-type zinc fingers.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: K02007; AAB59875.1; -!
 CC PIR: A03947; FOVNA2.
 CC HSSP: P05888; 1AAF.
 CC HIV: K02007; GAGSSF2.
 CC InterPro: IPR000721; Gag_p24.
 CC InterPro: IPR008916; Retrov_capsid_C.
 CC InterPro: IPR008919; Retrov_capsid_N.
 CC InterPro: IPR000071; Retrovir_p17.
 CC InterPro: IPR001878; Znf_CCHC.
 CC Pfam: PF00540; Gag_p17; 1.
 CC Pfam: PF00607; Gag_p24; 1.
 CC Pfam: PF00098; zf-CCHC; 2.
 CC PRINTS: PR00939; C2HCZNFINGER.
 CC PRINTS: PR00234; HIVMATRIX.
 DR

DR SMART: SM00343; ZNF C2HC; 2.
 DR PROSITE; PS50158; ZF CCHC; 2.
 KW AIDS; Core protein; Polypeptide; Myristate; Phosphorylation;
 ZNF-finger; Repeat; Lipoprotein.
 FT INIT MET 0
 FT CHAIN 1 133
 FT CHAIN 134 364
 FT CHAIN 365 378
 FT CHAIN 379 433
 FT CHAIN 434 449
 FT CHAIN 450 501
 FT ZN FING 391 408
 FT ZN FING 412 429
 FT LIPID 1
 FT SEQUENCE 501 AA; 55935 MW; 8F36C928D5EDBA45 CRC64;
 Query Match 35.7%; Score 50; DB 1; Length 501;
 Best Local Similarity 78.6%; Pred. No. 5.9;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 15 TERGRSLYNTVATL 28
 DB 71 SEELRSLYNTVATL 84
 RESULT 13
 ID GAG_HV1B1 STANDARD; PRT; 511 AA.
 AC P03347;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GAG polyprotein [Contains: Core proteins P17, P24, P2, P7, P1, P6].
 GN GAG.
 OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11678;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8511123; PubMed=2578615;
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
 RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
 RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
 RA Lautenberger J.A., Papas T.S., Ghayab J., Chang N.T., Gallo R.C.,
 RA Wong-Staal F.;
 RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
 RL Nature 313:277-284(1985).
 CC -!- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
 CC -!- PTM: The p24 protein is phosphorylated.
 CC -!- SIMILARITY: Contains 2 CCHC-type zinc fingers.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; M15654; AAA4201.1; -
 DR PIR; A03945; F0VWH3.
 DR HSP; P05888; IAAF.
 DR HIV; M15654; GAGSH102.
 DR InterPro; IPR000721; Gag_p24.
 DR InterPro; IPR008916; Retrov_capsid_C.
 DR InterPro; IPR008919; Retrov_capsid_N.

DR InterPro; IPR000721; Retrov p17.
 DR InterPro; IPR001878; Znf CCHC.
 DR Pfam; PF00540; Gag_p17; 1.
 DR Pfam; PF00607; Gag_p24; 1.
 DR PRINTS; PF00939; ZF-CCHC; 2.
 DR PRINTS; PF00234; HIVMATRIX.
 DR SMART; SM00343; ZNF C2HC; 2.
 DR PROSITE; PS50158; ZF CCHC; 2.
 KW AIDS; Core protein; Polypeptide; Myristate; Phosphorylation;
 ZNF-finger; Repeat; Lipoprotein.
 FT INIT MET 0
 FT CHAIN 1 131
 FT CHAIN 132 362
 FT CHAIN 363 376
 FT CHAIN 377 431
 FT CHAIN 432 447
 FT CHAIN 448 511
 FT ZN FING 389 406
 FT ZN FING 410 427
 FT LIPID 1
 FT SEQUENCE 511 AA; 57125 MW; 97C95188B80D458 CRC64;
 Query Match 35.7%; Score 50; DB 1; Length 511;
 Best Local Similarity 78.6%; Pred. No. 6;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 15 TERGRSLYNTVATL 28
 DB 71 SEELRSLYNTVATL 84
 RESULT 14
 ID GAG_HV1B5 STANDARD; PRT; 511 AA.
 AC P04593;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GAG polyprotein [Contains: Core proteins P17, P24, P2, P7, P1, P6].
 GN GAG.
 OS Human immunodeficiency virus type 1 (BHS isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11682;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8511123; PubMed=2578615;
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
 RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
 RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
 RA Lautenberger J.A., Papas T.S., Ghayab J., Chang N.T., Gallo R.C.,
 RA Wong-Staal F.;
 RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
 RL Nature 313:277-284(1985).
 CC -!- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
 CC -!- PTM: The p24 protein is phosphorylated.
 CC -!- SIMILARITY: Contains 2 CCHC-type zinc fingers.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```
DR EMBL; K02012; AAA44652.1; -.
DR HSSP; P05888; 1AAF.
DR HIV; K02012; GAGSBS.
DR InterPro; IPR000721; Gag_P24.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR000071; Retrovlt_p17.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00540; Gag_p17; 1.
DR Pfam; PF00607; Gag_P24; 1.
DR Pfam; PF00098; zf-CCHC; 2.
DR PRINTS; PR00939; C2HCZNFINGER.
DR PRINTS; PR00234; HIVMATRIX.
DR SMART; SM00343; Znf_C2HC; 2.
DR POSITE; PS50158; ZF_CCHC; 2.
KW AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat; Lipoprotein.
FT INIT_MET 0
FT CHAIN 1 131 BY SIMILARITY.
FT CHAIN 132 362 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 363 376 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 377 431 CORE PROTEIN P2.
FT CHAIN 432 447 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 448 511 CORE PROTEIN P1.
FT ZN_FING 389 406 CORE PROTEIN P6.
FT ZN_FING 410 427 CCHC-TYPE 1.
FT LPID 1 CCHC-TYPE 2.
SQ SEQUENCE 511 AA; 57107 MW; E0AE72D3CB8E7A80 CRC64;
Query Match 35.7%; Score 50; DB 1; Length 511;
Best local Similarity 78.6%; Pred. No. 6;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 15 TERGRSLYNTVATL 28
Db 71 SEELRSLYNTVATL 84

RESULT 15
GAG_HV1BR STANDARD; PRT; 511 AA.
ID GAG_HV1BR
AC P03348;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].
GN GAG.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OC NCB1_taxid=11686;
OX NCB1
RP SEQUENCE FROM N.A.
RX MEDLINE=85099333; PubMed=2981635;
RA Wain-Hobson S., Sonigo P., Dano O., Cole S., Alizon M.;
RT "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).
RN [2]
RP REVISIONS TO 459-470.
RX MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Wain-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients.";
RL Cell 46:63-74(1986).
CC -!- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -!- PTM: The p24 protein is phosphorylated.
CC -!- SIMILARITY: Contains 2 CCHC-type zinc fingers.
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02013; AAB59747.1; -.
DR HSSP; P05888; 1AAF.
DR HIV; K02013; GAGSBRU.
DR InterPro; IPR000721; Gag_P24.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR000071; Retrovlt_p17.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00540; Gag_p17; 1.
DR Pfam; PF00607; Gag_P24; 1.
DR Pfam; PF00098; zf-CCHC; 2.
DR PRINTS; PR00939; C2HCZNFINGER.
DR PRINTS; PR00234; HIVMATRIX.
DR SMART; SM00343; Znf_C2HC; 2.
DR POSITE; PS50158; ZF_CCHC; 2.
KW AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat; Lipoprotein.
FT INIT_MET 0
FT CHAIN 1 131 BY SIMILARITY.
FT CHAIN 132 362 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 363 376 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 377 431 CORE PROTEIN P2.
FT CHAIN 432 447 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 448 511 CORE PROTEIN P1.
FT ZN_FING 389 406 CORE PROTEIN P6.
FT ZN_FING 410 427 CCHC-TYPE 1.
FT LPID 1 CCHC-TYPE 2.
SQ SEQUENCE 511 AA; 57095 MW; 55972701A8434673 CRC64;
Query Match 35.7%; Score 50; DB 1; Length 511;
Best local Similarity 78.6%; Pred. No. 6;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 15 TERGRSLYNTVATL 28
Db 71 SEELRSLYNTVATL 84
```

Search completed: July 20, 2004, 06:31:00
Job time : 7.5082 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:43 ; Search time 28.2295 Seconds

(without alignments)
312.953 Million cell updates/sec

Title: US-09-673-166B-273

Perfect score: 140
Sequence: 1 GROYIKANSKFIGITERGRSLVNTVATL 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	50.0	1310	2 Q93N27	Q93N27 clostridium
2	55	39.3	342	16 Q98FZ0	Q98FZ0 rhizobium
3	55	39.3	499	15 Q8AU75	Q8AU75 human
4	54	38.6	80	15 Q9YK29	Q9YK29 human
5	54	38.6	82	15 Q9WEV6	Q9WEV6 human
6	54	38.6	82	15 Q9YKY7	Q9YKY7 human
7	54	38.6	101	15 Q9INB0	Q9INB0 human
8	54	38.6	102	15 Q9INB3	Q9INB3 human
9	54	38.6	103	15 Q9IND2	Q9IND2 human
10	54	38.6	119	15 Q9J661	Q9J661 human
11	54	38.6	121	15 Q99583	Q99583 human
12	54	38.6	121	15 Q99589	Q99589 human
13	54	38.6	124	15 Q99580	Q99580 human
14	54	38.6	126	15 Q99551	Q99551 human
15	54	38.6	126	15 Q995V6	Q995V6 human
16	54	38.6	126	15 Q995R8	Q995R8 human

17	54	38.6	127	15 Q77209	Q77209 human
18	54	38.6	128	15 Q8J6D4	Q8J6D4 human
19	54	38.6	129	15 Q55392	Q55392 human
20	54	38.6	129	15 Q78107	Q78107 human
21	54	38.6	130	15 Q8Q7Q9	Q8Q7Q9 human
22	54	38.6	130	15 Q76947	Q76947 human
23	54	38.6	130	15 Q78602	Q78602 human
24	54	38.6	130	15 Q76945	Q76945 human
25	54	38.6	130	15 Q77179	Q77179 human
26	54	38.6	130	15 Q78604	Q78604 human
27	54	38.6	130	15 Q77176	Q77176 human
28	54	38.6	130	15 Q76940	Q76940 human
29	54	38.6	130	15 Q8Q7T3	Q8Q7T3 human
30	54	38.6	130	15 Q78605	Q78605 human
31	54	38.6	130	15 Q77174	Q77174 human
32	54	38.6	130	15 Q8Q7Q8	Q8Q7Q8 human
33	54	38.6	130	15 Q78599	Q78599 human
34	54	38.6	130	15 Q78600	Q78600 human
35	54	38.6	130	15 Q78601	Q78601 human
36	54	38.6	130	15 Q77177	Q77177 human
37	54	38.6	130	15 Q76939	Q76939 human
38	54	38.6	130	15 Q55384	Q55384 human
39	54	38.6	130	15 Q78611	Q78611 human
40	54	38.6	130	15 Q8Q7S5	Q8Q7S5 human
41	54	38.6	130	15 Q78608	Q78608 human
42	54	38.6	130	15 Q78613	Q78613 human
43	54	38.6	130	15 Q76942	Q76942 human
44	54	38.6	130	15 Q76943	Q76943 human
45	54	38.6	130	15 Q77210	Q77210 human

ALIGNMENTS

RESULT 1

Q93N27 PRELIMINARY; PRT; 1310 AA.
ID Q93N27
AC Q93N27;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Tetanus toxin (Fragment).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RA Shunlin Z., Dianliang L.,
RT "Cloning and sequence analysis of tetanus toxin gene,"
RT Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF89424; AK72964.2; -
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0009405; P:patogenesis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008985; ConA-like_lcc_g1.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR002160; Kunitz_legume.
DR InterPro; IPR000395; Peptidase_M27.
DR InterPro; IPR006025; Pept M_zn_BS.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOTOXISIN.
DR PRODOM; PD001963; Bontotoxinin.1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
FT PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON_TER 1
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418F450 CRC64;
Query Match 50.0%; Score 70; DB 2; Length 1310;

Best Local Similarity 100.0%; Pred. No. 0.093;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 QYKANSKFIGITE 16
Db 831 QYKANSKFIGITE 844

RESULT 2
Q98FZ0 PRELIMINARY; PRT; 342 AA.
AC Q98FZ0;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein ml13560.
GN ML13560.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OX Phyllobacteriaceae; Mesorhizobium.
RN NCBI_TaxID=381;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=MAFP303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Matanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Yamada S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003002; BAB50426.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 342 AA; 37582 MW; 99597D57D25D11A1 CRC64;
QY
Query Match 39.3%; Score 55; DB 16; Length 342;
Best Local Similarity 41.7%; Pred. No. 4.6;
Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
Db 3 QYKANSKFIGITERGRSLYNTVA 26
82 RFLKAGSDFIGVADTGWFFCTVA 105
RESULT 3
Q8AU75 PRELIMINARY; PRT; 499 AA.
AC Q8AU75;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Gag protein.
GN Gag.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=99ZASW7;
RX MEDLINE=22190101; PubMed=12201911;
RA Papatianopoulos M.A., Cilliers T., Morris L., Mokili J.L.,
RA Dowling W., Bix D.L., McCutchan F.E.;
RT "Full-length Genome Analysis of HIV-1 Subtype C Utilizing CXCR4 and
Inter-subtype Recombinants Isolated in South Africa."
RL AIDS Res. Hum. Retroviruses 18:879-886(2002).
RN [2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=99ZASW7;
RA Papatianopoulos M.A., Morris L., Mokili J., Dowling W., Carr J.K.,
RA McCutchan F.;
RT Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL; AF411966; AAN47118.1; -
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016032; P:viral life cycle; IEA.
DR InterPro; IPR000721; Gag_p24.
DR InterPro; IPR000071; Retrovir_p17.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR001878; ZnF_CCHC.
DR Pfam; PF00540; Gag_p17; 1.
DR Pfam; PF00607; Gag_p24; 1.
DR Pfam; PF00098; zf-CCHC; 2.
DR PRINTS; PR00939; C2HCZPINGER.
DR PRINTS; PR00234; HIVMATRIX.
DR SMART; SM00343; ZNF_CCHC; 2.
DR PROSITE; PS50158; ZF_CCHC; 2.
SQ SEQUENCE 499 AA; 55941 MW; 1B4F1D83189BD712 CRC64;

Query Match 39.3%; Score 55; DB 15; Length 499;
Best Local Similarity 85.7%; Pred. No. 7;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
Db 72 TEERVSLYNTVATL 85

RESULT 4
Q9YKZ9 PRELIMINARY; PRT; 80 AA.
AC Q9YKZ9;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Gag protein (Gag polyprotein) (Fragment).
GN Gag.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=DG5218;
RX MEDLINE=99210134; PubMed=10195759;
RA Meghavi M.H., Salminen M.O., Sonnerborg A., Vahine A.;
RT "DNA sequence of the long terminal repeat of human immunodeficiency
virus type 1 subtype A through G."
RL AIDS Res. Hum. Retroviruses 15:485-488(1999).
DR EMBL; AF106178; AAD16923.1; -
DR GO; GO:0019012; C:viralion; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000071; Retrovir_p17.
DR Pfam; PF00540; Gag_p17; 1.
DR PRINTS; PR00234; HIVMATRIX.
KW AIDS; Core protein; Polyprotein.
FT NON TER 1
FT 80
FT 80
SQ SEQUENCE 80 AA; 9341 MW; E556DE35D7F1042F CRC64;
QY
Query Match 38.6%; Score 54; DB 15; Length 80;
Best Local Similarity 85.7%; Pred. No. 1.3;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 15 TERGRSLYNTVATL 28
46 TEERVSLYNTVATL 59
RESULT 5
Q9WEV6 PRELIMINARY; PRT; 82 AA.
AC Q9WEV6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Gag protein (Gag polypeptide) (Fragment).
GN GAG.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZMS710;
RX MEDLINE=99210134; PubMed=10195759;
RA Naghavi M.H., Salminen M.O., Sommerborg A., Vahne A.;
RT "DNA sequence of the long terminal repeat of human immunodeficiency
RT virus type 1 subtype A through G.";
RL AIDS Res. Hum. Retroviruses 15:485-488(1999).
DR EMBL: AF106192; AAD16937.1; -.
DR GO: GO:0019012; C:viral; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000071; Retroviral_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS: PR00234; HIVMATRIX.
KW AIDS: Core protein; Polypeptide.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 9229 MW; 7F8CDF198904E579 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 82;
Best Local Similarity 85.7%; Pred. No. 1.3;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
DB 48 TEELRSLYNTVATL 61

RESULT 6
Q9YKY7 PRELIMINARY; PRT; 82 AA.
ID Q9YKY7
AC Q9YKY7
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Gag protein (Gag polypeptide) (Fragment).
GN GAG.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SO6173;
RX MEDLINE=99210134; PubMed=10195759;
RA Naghavi M.H., Salminen M.O., Sommerborg A., Vahne A.;
RT "DNA sequence of the long terminal repeat of human immunodeficiency
RT virus type 1 subtype A through G.";
RL AIDS Res. Hum. Retroviruses 15:485-488(1999).
DR EMBL: AF106193; AAD16938.1; -.
DR GO: GO:0019012; C:viral; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000071; Retroviral_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS: PR00234; HIVMATRIX.
KW AIDS: Core protein; Polypeptide.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 9475 MW; 66A40A2AD4B4A30 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 82;
Best Local Similarity 85.7%; Pred. No. 1.3;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
DB 48 TEELRSLYNTVATL 61

RESULT 7
Q9INE0 PRELIMINARY; PRT; 101 AA.
ID Q9INE0
AC Q9INE0
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Gag protein (Gag polypeptide) (Fragment).
GN GAG.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DRCM24;
RX MEDLINE=99260287; PubMed=10331444;
RA Mokili J.L., Wade C.M., Burns S.M., Cutting W.A., Bopopi J.M.,
RA Green S.D., Peutherer J.F., Simmonds P.;
RT "Genetic heterogeneity of HIV type 1 subtypes in Kimpese, rural
RT Democratic Republic of Congo.";
RL AIDS Res. Hum. Retroviruses 15:655-664(1999).
DR EMBL: AF144809; AAF69035.1; -.
DR GO: GO:0019012; C:viral; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000071; Retroviral_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS: PR00234; HIVMATRIX.
KW AIDS: Core protein; Polypeptide.
FT NON_TER 1
FT NON_TER 101
SQ SEQUENCE 101 AA; 11256 MW; 3B601426CF7841C9 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 101;
Best Local Similarity 85.7%; Pred. No. 1.7;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
DB 33 TEELRSLYNTVATL 46

RESULT 8
Q9INA3 PRELIMINARY; PRT; 102 AA.
ID Q9INA3
AC Q9INA3
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Gag protein (Gag polypeptide) (Fragment).
GN GAG.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DRCM17;
RX MEDLINE=99260287; PubMed=10331444;
RA Mokili J.L., Wade C.M., Burns S.M., Cutting W.A., Bopopi J.M.,
RA Green S.D., Peutherer J.F., Simmonds P.;
RT "Genetic heterogeneity of HIV type 1 subtypes in Kimpese, rural
RT Democratic Republic of Congo.";
RL AIDS Res. Hum. Retroviruses 15:655-664(1999).
DR EMBL: AF144846; AAF69072.1; -.
DR GO: GO:0019012; C:viral; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000071; Retroviral_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS: PR00234; HIVMATRIX.
KW AIDS: Core protein; Polypeptide.
FT NON_TER 1
FT NON_TER 102
SQ SEQUENCE 102 AA; 11423 MW; 69A9761220982C5E CRC64;

Query Match 38.6%; Score 54; DB 15; Length 102;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 |||||
 DB 33 TEELRSLYNTVATL 46

RESULT 9
 Q9IND2 PRELIMINARY; PRT; 103 AA.

AC Q9IND2;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, last annotation update)
 DE Gag protein (Gag polypeptide) (Fragment).
 GN GAG.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 NC NCBI_TaxID=11676;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BXOM13;
 RX MEDLINE=9260287; PubMed=10311444;
 RA McKill J.L., Wade C.M., Burns S.M., Cutting W.A., Bopopi J.M.,
 Green S.D., Penhagher J.F., Simmonds P.;
 RT "Genetic heterogeneity of HIV type 1 subtypes in Kimpese, rural
 Democratic Republic of Congo."
 RL AIDS Res. Hum. Retroviruses 15:655-664(1999).
 DR EMBL; AF144817; AAF69043.1; -
 DR GO; GO:0019012; C:virion; IEA.
 DR InterPro; IPR000071; Retroviral_p17.
 DR PRINTS; PR00540; Gag_p17; 1
 DR AIDS; Core protein; Polypeptide.
 KW NON_TER 1
 FT NON_TER 103
 SQ SEQUENCE 103 AA; 11577 MW; AAD469474ED731E2 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 |||||
 DB 33 TEELRSLYNTVATL 46

RESULT 10
 Q9J661 PRELIMINARY; PRT; 119 AA.

AC Q9J661;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, last annotation update)
 DE Gag protein (Gag polypeptide) (Fragment).
 GN GAG.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 NC NCBI_TaxID=11676;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BX6AIP41;
 RX Iversen A.K., Learn G.H., Fugger L., Gerstoft J., Mullins J.I.,
 Skirboj P.;
 RT "Presence of multiple HIV subtypes and a high frequency of subtype
 RT chimeric viruses in heterosexually infected women."
 RL J. Acquir. Immune Defic. Syndr. Hum. Retroviral. 22:325-332(1999).
 DR EMBL; AF191484; AAF37550.1; -
 DR GO; GO:0019012; C:virion; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000071; Retroviral_p17.
 DR Pfam; PF00540; Gag_p17; 1
 DR PRINTS; PR00234; HIVMATRIX.
 KW AIDS; Core protein; Polypeptide.
 FT NON_TER 1
 FT NON_TER 119
 SQ SEQUENCE 119 AA; 13536 MW; F1366864F6F28DD9 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 119;
 Best Local Similarity 85.7%; Pred. No. 2;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 |||||
 DB 42 TEELRSLYNTVATL 55

RESULT 11
 Q995R9 PRELIMINARY; PRT; 121 AA.

AC Q995R9;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, last annotation update)
 DE Gag protein (Gag polypeptide) (Fragment).
 GN GAG.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 NC NCBI_TaxID=11676;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Y9;
 RX Vitrelli D.L.;
 RT "Immunological and virological characteristics of HIV infected
 RT individuals in different risk groups in Madras, India."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF250270; AAK34984.1; -
 DR GO; GO:0019012; C:virion; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000071; Retroviral_p17.
 DR Pfam; PF00540; Gag_p17; 1
 DR PRINTS; PR00234; HIVMATRIX.
 KW AIDS; Core protein; Polypeptide.
 FT NON_TER 1
 FT NON_TER 121
 SQ SEQUENCE 121 AA; 13968 MW; 369B37CDA213AD4F CRC64;

Query Match 38.6%; Score 54; DB 15; Length 121;
 Best Local Similarity 85.7%; Pred. No. 2.1;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 |||||
 DB 48 TEELRSLYNTVATL 61

RESULT 12
 Q41399 PRELIMINARY; PRT; 121 AA.

AC Q41399;
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-JAN-1998 (TRENBLrel. 05, last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, last annotation update)
 DE P17 (Gag polypeptide) (Fragment).
 GN GAG.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 NC NCBI_TaxID=11676;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ET2555;
 RX Shereta K., Johansson B., Sommerborg A.;


```
RT "distribution of HIV-1 subtypes in Ethiopia."
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U56345; AB67649.1; -.
DR GO: GO:0019012; C:vitation; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS: PR00234; HIVMATRIX.
KM AIDS; Core protein; Polyprotein.
FT NON_TER 121 121
SQ SEQUENCE 121 AA; 13849 MW; 89661A7FE09CD1B3 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 121;
Best Local Similarity 85.7%; Pred. No. 2.1;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
DB 72 TEELRSLYNTVATL 85

RESULT 13
Q995S0 PRELIMINARY; PRT; 124 AA.
AC Q995S0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Gag protein (Gag polyprotein) (Fragment).
GN GAG.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1] _TaxID=11676;
RP SEQUENCE FROM N.A.
RC STRAIN=Y8;
RA Yirell D.L.;
RT "Immunological and virological characteristics of HIV infected individuals in different risk groups in Madras, India."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF250269; AAK34983.1; -.
DR GO: GO:0019012; C:vitation; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS: PR00234; HIVMATRIX.
KM AIDS; Core protein; Polyprotein.
FT NON_TER 1 1
FT NON_TER 124 124
SQ SEQUENCE 124 AA; 14178 MW; C5B1B05020568A77 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 124;
Best Local Similarity 85.7%; Pred. No. 2.1;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
DB 46 TEELRSLYNTVATL 59

RESULT 14
Q995S1 PRELIMINARY; PRT; 126 AA.
AC Q995S1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Gag protein (Gag polyprotein) (Fragment).
GN GAG.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=Y7;
RA Yirell D.L.;
RT "Immunological and virological characteristics of HIV infected individuals in different risk groups in Madras, India."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF250268; AAK34982.1; -.
DR GO: GO:0019012; C:vitation; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS: PR00234; HIVMATRIX.
KM AIDS; Core protein; Polyprotein.
FT NON_TER 1 1
FT NON_TER 126 126
SQ SEQUENCE 126 AA; 14336 MW; 5C561821DBED0D4 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 126;
Best Local Similarity 85.7%; Pred. No. 2.2;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
DB 48 TEELRSLYNTVATL 61

RESULT 15
Q99FV6 PRELIMINARY; PRT; 126 AA.
AC Q99FV6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Gag protein (Gag polyprotein) (Fragment).
GN GAG.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B197;
RA MEDLINE=99181514; PubMed=10081673;
RA Ramos A., Tamuri A., Schechter M., Rayfield M.A., Hu D.J.,
RA Cabral M.C., Bandea C.I., Baggis J., Pieniazek D.,
RT "Dual and recombinant infections: an integral part of the HIV-1 epidemic in Brazil."
RL Emerging Infect. Dis. 5:65-74(1999).
DR EMBL: AF115451; AAD21596.1; -.
DR GO: GO:0019012; C:vitation; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS: PR00234; HIVMATRIX.
KM AIDS; Core protein; Polyprotein.
FT NON_TER 1 1
FT NON_TER 126 126
SQ SEQUENCE 126 AA; 14137 MW; 6EC9B47F2FD3CF80 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 126;
Best Local Similarity 85.7%; Pred. No. 2.2;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
DB 72 TEELRSLYNTVATL 85
```

Search completed: July 20, 2004, 06:33:15
Job time : 30.2295 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:07:17 ; Search time 41.7705 Seconds

(without alignments)
189,400 Million cell updates/sec

Title: US-09-673-166b-273

Perfect score: 140

Sequence: 1 GROYIKANKSKFIFGITERGRSLYNTVATL 28

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	70.0	28	2	AAV53552
2	88.5	63.2	25	4	AA844738
3	87	62.1	29	2	AAV53551
4	83	59.3	29	2	AAW48991
5	83	59.3	29	2	AAW48992
6	83	59.3	29	2	AAW09326
7	83	59.3	29	2	AAW09327
8	83	59.3	29	3	AA812418
9	83	59.3	29	3	AA812419
10	83	59.3	29	4	AAE05438
11	83	59.3	29	4	AAE05437
12	83	59.3	29	5	AAU10851
13	83	59.3	29	5	AAU10850
14	82	58.6	28	5	AAU11422
15	81	57.9	29	5	AA879182
16	80	57.1	158	5	AAW81330
17	80	57.1	158	5	AA879182
18	79	56.4	218	5	AAV49275
19	79	56.4	240	3	AAV49254
20	79	56.4	390	3	AAV49255
21	79	56.4	514	6	AAO30491
22	79	56.4	514	6	AAO30490
23	79	56.4	514	6	AAO30495
24	79	56.4	517	6	AAO30492
25	77	55.0	750	3	AAV92645

26	77	55.0	750	3	AAV92638	AAV92638 Mutant hu
27	77	55.0	750	3	AAV92640	AAV92640 Mutant hu
28	76	54.3	25	3	AAV92651	AAV92651 psmep008
29	76	54.3	29	2	AA83561	AA83561 Ige CH4 I
30	76	54.3	37	2	AA865389	AA865389 Universal
31	76	54.3	37	2	AA865383	AA865383 Universal
32	76	54.3	47	2	AA862723	AA862723 LHRH-cont
33	76	54.3	750	3	AAV92627	AAV92627 Mutant hu
34	75.5	53.9	750	3	AAV92639	AAV92639 Mutant hu
35	75	53.6	25	3	AAV92652	AAV92652 psmep009
36	75	53.6	50	2	AAW06132	AAW06132 Anti-choi
37	75	53.6	50	2	AAW46447	AAW46447 CERP B ce
38	75	53.6	109	4	AA820145	AA820145 Growth di
39	75	53.6	124	3	AA845492	AA845492 Modified
40	75	53.6	124	3	AA845517	AA845517 Modified
41	75	53.6	158	2	AAW81329	AAW81329 TNF2-4, a
42	75	53.6	158	5	AA807281	AA807281 Human TNF
43	75	53.6	216	3	AAV92665	AAV92665 MGC-1 ana
44	75	53.6	693	3	AAV92649	AAV92649 Mutant hu
45	75	53.6	750	3	AAV92630	AAV92630 Mutant hu

ALIGNMENTS

RESULT 1	
AAV53552	
ID	AAV53552 standard; protein; 28 AA.
XX	
AC	AAV53552;
XX	
DT	18-JAN-2000 (first entry)
XX	
DE	Lipopeptide #3.
XX	
KW	Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
KW	electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
KW	human immunodeficiency virus; hepatitis B virus; papilloma virus;
XX	melanoma; malaria; parasite.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
FH	Key
FT	Modified-site
FT	1
FT	/note= "contains palmitoyl residue attached at the N-
XX	terminus and on the epsilon carbon of the side chain"
XX	
FN	FR2776926-A1.
XX	
PD	08-OCT-1999.
XX	
PF	07-APR-1998; 98FR-00004323.
XX	
PR	07-APR-1998; 98FR-00004323.
XX	
PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA	(CNRS) CNRS CENT NAT RECH SCI.
PA	(INSP) INST PASTEUR ILLRE.
XX	
PI	Le Gal FA, Guillet JG, Gahery SH, Gras MH, Melnyk O, Tartar A;
XX	
DR	WPI; 1999-583113/50.
XX	
PT	New lipopeptide containing lipid regions and two epitopes, all separated
PT	by peptide spacers that impart hydrophilicity, useful in vaccines.
XX	
PS	Example 1; Page 9; 35pp; French.
XX	
CC	The invention relates to the generation of a lipopeptide comprising at
CC	least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)
CC	epitope and at least one lipid residue with (i) the epitopes and lipid
CC	portion and (ii) the epitopes, being separated independently by peptide

CC spacers. These spacers comprise sequences of amino acids which carry an
CC overall electrical charge in neutral media to ensure that the lipopeptide
CC is hydrophilic. This peptide represents an example of them lipopeptide of
CC the invention. It contains 2 lipid residues attached at the N-terminal
CC end of the molecule and 2 epitopes separated by the spacer residues Ser-
CC Ser and Ala-Ala-Ala. Peptides AA53301-Y53549 represent peptide epitopes
CC used in the generation of the lipopeptides. These are used in therapeutic
CC or prophylactic compositions and vaccines to induce specific immune
CC responses against human immunodeficiency, hepatitis B or papilloma
CC viruses; p53 of melanoma or the malaria parasite

SQ Sequence 28 AA;

Query Match 70.0%; Score 98; DB 2; Length 28;
Best Local Similarity 95.0%; Pred. No. 1.5e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GROVIKANSKFIGITERGRS 20
1 GROVIKANSKFIGITERGRA 20

RESULT 2

ID AAB84738 standard; peptide; 25 AA.
AC AAB84738;
DT 17-SEP-2001 (first entry)

DE Amino acid sequence of lipopeptide Nh2-K.

KW Lipid-tailed protein; mucosal membrane; immune system; lipoprotein;
KW B cell response; T cell response; intranasal immunisation;
KW sublingual immunisation.

OS Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1
FT /note= "amidated residue with palmitoyl lipid residue
FT with 16 carbon atoms attached via an epsilon bond"

PN WO200141797-A2.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000WO-EP012794.

PR 09-DEC-1999; 99US-0169952P.

PA (INSP) INST PASTEUR.

PA (UYLI-) UNIV LILLE 2.

PI Druilhe P, Gras-Masse H, Benmohamed L,

DR WPI; 2001-451597/48.

PT Inducing systemic immune response comprises mucosal administration of

PS Claim 22; Page 26; 34pp; English.

CC The present sequence represents a lipid-tailed protein. The lipid moiety
CC is able to cross the mucosal membrane, and deliver an antigen to the
CC immune system. The lipoprotein is administered to a mucosal membrane to
CC induce an immune response. The lipoprotein induces a B cell and/or a T
CC cell response. The lipoprotein is used to provide intranasal or
CC sublingual immunization

SQ Sequence 25 AA;

Query Match 63.2%; Score 88.5; DB 4; Length 25;

Best Local Similarity 90.9%; Pred. No. 4.5e-07;
Matches 20; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 GROVI-KANSKFIGITERGRSL 21
1 GROVIKANSKFIGITERGRIL 22

RESULT 3
ID AAY53551 standard; protein; 29 AA.

AC AAY53551;

DT 18-JAN-2000 (first entry)

DE Lipopeptide #2.

KW Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
KW electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
KW human immunodeficiency virus; hepatitis B virus; papilloma virus;
KW melanoma; malaria; parasite.

OS Synthetic.
OS Homo sapiens.

FH Key Location/Qualifiers

FT Modified-site 1
FT /note= "contains palmitoyl residue attached at the N-
FT terminus and on the epsilon carbon of the side chain"

PN FR2776926-A1.

PD 08-OCT-1999.

PF 07-APR-1998; 98FR-00004323.

PR 07-APR-1998; 98FR-00004323.

PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PA (CNRS) CNRS CENT NAT RECH SCI.

PA (INSP) INST PASTEUR LILLE.

PI Le Gal FA, Guillet JG, Gahery SH, Gras MH, Melnyk O, Tartar A;

DR WPI; 1999-583113/50.

PT New lipopeptide containing lipid regions and two epitopes, all separated

PS Example 1; Page 9; 35pp; French.

CC The invention relates to the generation of a lipopeptide comprising at
CC least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)
CC epitope and at least one lipid residue with (i) the epitopes and lipid
CC portion and (ii) the epitopes, being separated independently by peptide
CC spacers. These spacers comprise sequences of amino acids which carry an
CC overall electrical charge in neutral media to ensure that the lipopeptide
CC is hydrophilic. This peptide represents an example of them lipopeptide of
CC the invention. It contains 2 lipid residues attached at the N-terminal
CC end of the molecule and 2 epitopes separated by the spacer residues Ser-
CC Ser and Ala-Ala-Ala. Peptides AA53301-Y53549 represent peptide epitopes
CC used in the generation of the lipopeptides. These are used in therapeutic
CC or prophylactic compositions and vaccines to induce specific immune
CC responses against human immunodeficiency, hepatitis B or papilloma
CC viruses; p53 of melanoma or the malaria parasite

SQ Sequence 29 AA;

Query Match 62.1%; Score 87; DB 2; Length 29;
Best Local Similarity 94.4%; Pred. No. 9.4e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 OYIKANSKFIGTTERGRS 20
|||||
Db 4 OYIKANSKFIGTTERGRA 21

RESULT 4
AAW48991

ID AAW48991 standard; peptide; 29 AA.

AC AAW48991;

DT 23-SEP-1998 (first entry)

DE Lipidated vaccine 1 against human cytomegalovirus infection.

KW Immunogenic epitope; human cytomegalovirus; HCMV; vaccine; CD8+;
KW cytotoxic T-lymphocyte; MHC; major histocompatibility complex;
KW helper T-lymphocyte; HTL; CTL; palmitic acid.

OS Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1 /note= "attached to two palmitic acid molecules"

FT Region 2..17 /note= "Helper T-cell epitope"

FT Region 18..20 /note= "Spacer"

FT Region 21..29

FT /note= "Cytotoxic T-cell epitope referred to as the 495
FT peptide"

PN W09821233-A2.

PD 22-MAY-1998.

PF 12-NOV-1997; 97WO-US020236.

PR 12-NOV-1996; 96US-00747488.

PR 14-OCT-1997; 97US-00950064.

XX (CITY) CITY OF HOPE.

PA Diamond DJ, York U;

PI WPI; 1998-297862/26.

PT Immunoreactive human cytomegalovirus epitopes - useful to vaccinate
PT against infection.

PS Example 11; Page 31; 56pp; English.

CC The present lipidated vaccine comprises of two palmitic acid molecules
CC attached to a peptide portion. The peptide portion contains a helper T-
CC lymphocyte (HTL) epitope attached to a spacer which in turn is attached
CC to a cytotoxic T-lymphocyte (CTL) epitope. The CTL epitope is referred to
CC as the 495 peptide as it represents residues 495-503 of pp65 from human
CC cytomegalovirus (HCMV) laboratory strains AD169 and Towne. The invention
CC provides peptides which are immunogenic epitopes that are recognised by
CC the CD8+ class I major histocompatibility complex (MHC) restricted
CC cytotoxic T-lymphocytes of patients harboring latent HCMV infection. The
CC peptides are claimed to be capable of activating cytotoxic T-cell
CC lymphocytes in the absence of active viral replication, and thus are
CC useful for eliciting a cellular immune response against HCMV by normal
CC and immunodeficient subjects. The immunological peptides can also be
CC formulated as vaccines, such as the present one, which are claimed to be
CC useful for protecting against HCMV infection, augmenting the immune
CC system response to a HCMV infection or protecting against reactivation of
CC a latent HCMV infection

XX Sequence 29 AA;

Query Match 59.3%; Score 83; DB 2; Length 29;

Best Local Similarity 69.2%; Pred. No. 4.2e-06;
Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 OYIKANSKFIGTTERGRSIXNTATL 28
|||||
Db 4 OYIKANSKFIGTTEAANLVPMVATV 29

RESULT 5
AAW48992

ID AAW48992 standard; peptide; 29 AA.

AC AAW48992;

DT 23-SEP-1998 (first entry)

DE Lipidated vaccine 2 against human cytomegalovirus infection.

KW Immunogenic epitope; human cytomegalovirus; HCMV; vaccine; CD8+;
KW cytotoxic T-lymphocyte; MHC; major histocompatibility complex;
KW helper T-lymphocyte; HTL; CTL; palmitic acid.

OS Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1 /note= "attached to three palmitic acid molecules"

FT Region 4..17 /note= "Helper T-cell epitope"

FT Region 18..20 /note= "Spacer"

FT Region 21..29

FT /note= "Cytotoxic T-cell epitope referred to as the 495
FT peptide"

PN W09821233-A2.

PD 22-MAY-1998.

PF 12-NOV-1997; 97WO-US020236.

PR 12-NOV-1996; 96US-00747488.

PR 14-OCT-1997; 97US-00950064.

XX (CITY) CITY OF HOPE.

PA Diamond DJ, York U;

PI WPI; 1998-297862/26.

PT Immunoreactive human cytomegalovirus epitopes - useful to vaccinate
PT against infection.

PS Example 11; Page 31; 56pp; English.

CC The present lipidated vaccine comprises of three palmitic acid molecules
CC attached to a peptide portion. The peptide portion contains a helper T-
CC lymphocyte (HTL) epitope attached to a spacer which in turn is attached
CC to a cytotoxic T-lymphocyte (CTL) epitope. The CTL epitope is referred to
CC as the 495 peptide as it represents residues 495-503 of pp65 from human
CC cytomegalovirus (HCMV) laboratory strains AD169 and Towne. The invention
CC provides peptides which are immunogenic epitopes that are recognised by
CC the CD8+ class I major histocompatibility complex (MHC) restricted
CC cytotoxic T-lymphocytes of patients harboring latent HCMV infection. The
CC peptides are claimed to be capable of activating cytotoxic T-cell
CC lymphocytes in the absence of active viral replication, and thus are
CC useful for eliciting a cellular immune response against HCMV by normal
CC and immunodeficient subjects. The immunological peptides can also be
CC formulated as vaccines, such as the present one, which are claimed to be
CC useful for protecting against HCMV infection, augmenting the immune
CC system response to a HCMV infection or protecting against reactivation of
CC a latent HCMV infection

XX Sequence 29 AA;

Sequence 29 AA; 59.3%; Score 83; DB 2; Length 29;
Query Match Best Local Similarity 69.2%; Pred. No. 4.2e-06;
Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QYIKANSKFIIGTERGRSLYNTVATL 28
DB 4 QYIKANSKFIIGTEAANLVPMVATV 29

RESULT 6
ID AAY09326 standard; peptide; 29 AA.
XX AAY09326;
XX
XX
DT 08-JUL-1999 (first entry)
XX
DE Lipidated vaccine peptide SEQ ID NO:13.
XX
XX Human cytomegalovirus; hCMV; immunologically active peptide; vaccine;
KM immune response; cytotoxic T lymphocyte; CTL; immunostimulation;
KM infection; immunosuppression; bone marrow transplant; solid organ; heart;
KM AIDS.
XX
XX Synthetic.
OS
XX WO919349-A1.
XX
XX 22-APR-1999.
XX
PD 11-MAY-1998; 98WO-US009652.
PF
XX 14-OCT-1997; 97US-00950064.
PR 10-FEB-1998; 98US-00021298.
XX
XX (CITY) CITY OF HOPE.
PI
XX Diamond DJ, York J;
PI
DR MPI; 1999-277590/23.
XX
XX Immunogenic peptide cytotoxic T lymphocyte epitopes of human
PT cytomegalovirus.
XX
XX Example 11; Page 32; 64pp; English.

The present invention describes immunologically active peptides (IAPs) capable of eliciting a cellular immune response to human cytomegalovirus (hCMV). The IAP can be used in a (cellular) vaccine to augment the immune system response to hCMV, or to provide immunity against hCMV. The IAP (cellular) vaccine can also protect an individual having a latent hCMV infection from reactivation. An antigen presenting cell can be used to determine the presence or absence of hCMV-infected T lymphocytes. The viral vector containing IAP encoding DNA can also be used to provide immunity against hCMV. The IAP can be used to prepare hCMV-reactive human cytotoxic T lymphocytes. The antigen presenting cells primed with the IAPs can be used as diagnostic reagents to detect immunostimulation by hCMV. They can also detect active hCMV infection or exposure to hCMV. hCMV can cause opportunistic infections resulting in a variety of complications in, e.g. immunosuppressed patients. The IAP vaccines impart immunity to bone marrow transplant recipients, solid organ recipients, heart patients, AIDS patients or women of child-bearing years, without the need for ex vivo expansion of hCMV-specific cytotoxic T lymphocytes (which requires sophisticated laboratory setup and is highly labor-intensive and costly). The present sequence represents a lipidated vaccine peptide from the present invention

Sequence 29 AA;

Query Match 59.3%; Score 83; DB 2; Length 29;
Best Local Similarity 69.2%; Pred. No. 4.2e-06;

Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QYIKANSKFIIGTERGRSLYNTVATL 28
DB 4 QYIKANSKFIIGTEAANLVPMVATV 29

RESULT 7
ID AAY09327 standard; peptide; 29 AA.
XX AAY09327;
XX
XX
DT 08-JUL-1999 (first entry)
XX
DE Lipidated vaccine peptide SEQ ID NO:14.
XX
XX Human cytomegalovirus; hCMV; immunologically active peptide; vaccine;
KM immune response; cytotoxic T lymphocyte; CTL; immunostimulation;
KM infection; immunosuppression; bone marrow transplant; solid organ; heart;
KM AIDS.
XX
XX Synthetic.
OS
XX WO919349-A1.
XX
XX 22-APR-1999.
PF
XX 11-MAY-1998; 98WO-US009652.
PR 14-OCT-1997; 97US-00950064.
PR 10-FEB-1998; 98US-00021298.
XX
XX (CITY) CITY OF HOPE.
PI
XX Diamond DJ, York J;
PI
DR MPI; 1999-277590/23.
XX
XX Immunogenic peptide cytotoxic T lymphocyte epitopes of human
PT cytomegalovirus.
XX
XX Example 11; Page 32; 64pp; English.

The present invention describes immunologically active peptides (IAPs) capable of eliciting a cellular immune response to human cytomegalovirus (hCMV). The IAP can be used in a (cellular) vaccine to augment the immune system response to hCMV, or to provide immunity against hCMV. The IAP (cellular) vaccine can also protect an individual having a latent hCMV infection from reactivation. An antigen presenting cell can be used to determine the presence or absence of hCMV-infected T lymphocytes. The viral vector containing IAP encoding DNA can also be used to provide immunity against hCMV. The IAP can be used to prepare hCMV-reactive human cytotoxic T lymphocytes. The antigen presenting cells primed with the IAPs can be used as diagnostic reagents to detect immunostimulation by hCMV. They can also detect active hCMV infection or exposure to hCMV. hCMV can cause opportunistic infections resulting in a variety of complications in, e.g. immunosuppressed patients. The IAP vaccines impart immunity to bone marrow transplant recipients, solid organ recipients, heart patients, AIDS patients or women of child-bearing years, without the need for ex vivo expansion of hCMV-specific cytotoxic T lymphocytes (which requires sophisticated laboratory setup and is highly labor-intensive and costly). The present sequence represents a lipidated vaccine peptide from the present invention

Sequence 29 AA;

Query Match 59.3%; Score 83; DB 2; Length 29;
Best Local Similarity 69.2%; Pred. No. 4.2e-06;

DB 4 QYIKANSKFIGITEAANLVPMVATV 29

RESULT 8

AAB12418

ID AAB12418 standard; peptide; 29 AA.

AC AAB12418;

DT 20-OCT-2000 (first entry)

DE HCMV lipidated vaccine peptide SEQ ID NO:13.

XX Immunogenic; cytotoxic T lymphocytes; CTL; cytomegalovirus; HCMV;
XX infection; immune response; vaccine; immunostimulant; antiviral;
XX immunosuppressive; immunity; immunisation.

OS Human herpesvirus 5.

FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminally attached to (Pam)2 where Pam is
FT palmitic acid"

XX US6074645-A.

XX 13-JUN-2000.

XX 11-MAY-1998; 98US-00075257.

XX 12-NOV-1996; 96US-00747488.

XX 14-OCT-1997; 97US-00950064.

XX 10-FEB-1998; 98US-00021298.

XX (CITY) CITY OF HOPE.

XX York J, Diamond DJ;

XX WPI; 2000-430383/37.

XX New immunogenic peptides useful for providing immunity against
XX cytomegalovirus (CMV) infections, are capable of eliciting cellular
XX immune response to human CMV.

XX Example 11; Col 15; 18pp; English.

XX The present invention describes a cellular vaccine (I), which elicits a
XX MHC Class I cellular immune response to human cytomegalovirus (HCMV), and
XX comprises a peptide selected from (i) - (v), provided that (i) is not
XX another peptide (vi): (i) Asn Xaa1 Val Pro Met Val Ala Thr Xaa2; (ii) Tyr
XX Ala Leu Xaa5; (iv) Thr Pro Arg Val Thr Gly Gly Ala Xaa6; and (v) Phe
XX Pro Thr Lys Asp Val Ala Leu, where: Xaa1 = Ser, Thr or Leu; Xaa4 = Val
XX or Thr; Xaa5 = Leu, Arg or Lys; Xaa6 = Leu, Met or Phe; and (i) is not
XX Asn Leu Val Pro Met Val Ala Thr Val (vi). The peptides are useful for
XX providing immunity against future infections by HCMV, and for augmenting
XX the immune response of an individual who is latently infected with HCMV
XX and is at risk for reactivation of HCMV infection. The peptides are also
XX useful for imparting immunity to a bone marrow transplant recipient, a
XX solid organ recipient, a heart patient, an AIDS patient or a woman of
XX child-bearing years, without the need for ex vivo expansion of HCMV-
XX specific CTL. These peptides can directly stimulate cytotoxic T
XX lymphocytes (CTL) in vitro, thus can be used in an assay to determine the
XX degree of immunostimulation caused by HCMV. These may also be used to
XX distinguish individuals who are seropositive from those who have not been
XX exposed to HCMV and in the study of the Class I antigen-processing
XX pathway for HCMV proteins. The present sequence represents a peptide used
XX in the exemplification of the present invention

SQ Sequence 29 AA;

Query Match

59.3%; Score 83; DB 3; Length 29;

Best Local Similarity 69.2%; Pred. No. 4.2e-06;
Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QYIKANSKFIGITEGRSLYNTVATL 28

DB 4 QYIKANSKFIGITEAANLVPMVATV 29

RESULT 9

AAB12419

ID AAB12419 standard; peptide; 29 AA.

AC AAB12419;

DT 20-OCT-2000 (first entry)

DE HCMV lipidated vaccine peptide SEQ ID NO:14.

XX Immunogenic; cytotoxic T lymphocytes; CTL; cytomegalovirus; HCMV;
XX infection; immune response; vaccine; immunostimulant; antiviral;
XX immunosuppressive; immunity; immunisation.

OS Human herpesvirus 5.

FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminally attached to (Pam)3 where Pam is
FT palmitic acid"

XX US6074645-A.

XX 13-JUN-2000.

XX 11-MAY-1998; 98US-00075257.

XX 12-NOV-1996; 96US-00747488.

XX 14-OCT-1997; 97US-00950064.

XX 10-FEB-1998; 98US-00021298.

XX (CITY) CITY OF HOPE.

XX York J, Diamond DJ;

XX WPI; 2000-430383/37.

XX New immunogenic peptides useful for providing immunity against
XX cytomegalovirus (CMV) infections, are capable of eliciting cellular
XX immune response to human CMV.

XX Example 11; Col 15; 18pp; English.

XX The present invention describes a cellular vaccine (I), which elicits a
XX MHC Class I cellular immune response to human cytomegalovirus (HCMV), and
XX comprises a peptide selected from (i) - (v), provided that (i) is not
XX another peptide (vi): (i) Asn Xaa1 Val Pro Met Val Ala Thr Xaa2; (ii) Tyr
XX Ala Leu Xaa5; (iv) Thr Pro Arg Val Thr Gly Gly Ala Xaa6; and (v) Phe
XX Pro Thr Lys Asp Val Ala Leu, where: Xaa1 = Ser, Thr or Leu; Xaa4 = Val
XX or Thr; Xaa5 = Leu, Arg or Lys; Xaa6 = Leu, Met or Phe; and (i) is not
XX Asn Leu Val Pro Met Val Ala Thr Val (vi). The peptides are useful for
XX providing immunity against future infections by HCMV, and for augmenting
XX the immune response of an individual who is latently infected with HCMV
XX and is at risk for reactivation of HCMV infection. The peptides are also
XX useful for imparting immunity to a bone marrow transplant recipient, a
XX solid organ recipient, a heart patient, an AIDS patient or a woman of
XX child-bearing years, without the need for ex vivo expansion of HCMV-
XX specific CTL. These peptides can directly stimulate cytotoxic T
XX lymphocytes (CTL) in vitro, thus can be used in an assay to determine the
XX degree of immunostimulation caused by HCMV. These may also be used to
XX distinguish individuals who are seropositive from those who have not been
XX exposed to HCMV and in the study of the Class I antigen-processing
XX pathway for HCMV proteins. The present sequence represents a peptide used

CC	in the exemplification of the present invention
XX	
SQ	Sequence 29 AA;
XX	
XX	Query Match 59.3%; Score 83; DB 3; Length 29;
XX	Best local Similarity 69.2%; Pred. No. 4.2e-06;
XX	Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY	3 QYTKANSKFIGITERGRLVYVATL 28
DB	4 QYTKANSKFIGITEAANLVYVATV 29
XX	
XX	RESULT 10
XX	AAE05438
XX	ID AAE05438 standard; peptide: 29 AA.
XX	AC AAE05438;
XX	24-SEP-2001 (first entry)
DT	
XX	
DE	Human cytomegalovirus (HCMV) vaccine peptide #2.
XX	
XX	Human cytomegalovirus; HCMV; infection; vaccine; cytotoxic T lymphocyte;
KM	CTL; CTL precursor; CTLp; Class I antigen;
KW	major histocompatibility complex; MHC; cellular immune response;
KW	CTL activator; CTLp activator; immunostimulant.
XX	
OS	Human cytomegalovirus.
XX	
XX	Key Location/Qualifiers
FT	1
FT	/note= "Cysteine modified with palmitic acid"
FT	4..17
FT	/label= HTL_epitope
FT	18..20
FT	/note= "Amino acid spacer"
FT	21..29
FT	/label= CTL_epitope
XX	
PN	US6251399-B1.
XX	
PD	26-JUN-2001.
XX	
PF	27-MAR-2000; 2000US-00534639.
XX	
PR	12-NOV-1996; 96US-00747488.
PR	14-OCT-1997; 97US-00950064.
PR	10-FEB-1998; 98US-00021298.
PR	11-MAY-1998; 98US-00075257.
XX	
PA	(DIAM/) DIAMOND D J.
PA	(YORK/) YORK J.
PI	
PI	Diamond DJ, York J;
XX	
XX	WPI, 2001-431950/46.
DR	
XX	
XX	New peptides that are immunogenic epitopes of the human cytomegalovirus
PT	(CMV), useful for activating cytotoxic T lymphocytes (CTL) or CTL
PT	precursors to elicit an immune response against human CMV by normal or
PT	immunodeficient subjects.
XX	
PS	Example 11; Col 16; 17pp; English.
XX	
XX	The invention relates to immunologically active peptides, and functional
CC	variants thereof, capable of eliciting a cellular immune response to
CC	human cytomegalovirus (HCMV) in humans. The peptides are capable of
CC	directing human cytotoxic T lymphocytes (CTL) to recognise and lyse human
CC	cells infected with HCMV. Such immunologically active peptides, in
CC	association with an major histocompatibility complex (MHC) Class I
CC	molecule, are recognised by CTLs of individuals having a latent
CC	(inactive) HCMV infection. Vaccines comprising these peptides are useful

CC	for activating CTLs and CTL precursors (CTLP), particularly for eliciting
CC	a cellular immune response against HCMV by normal and immunodeficient
CC	subjects. The peptides may be used in an assay for determining the degree
CC	of immunostimulation caused by HCMV. The peptides may also be used for
CC	distinguishing individuals who are seropositive from those who have not
CC	been exposed to HCMV (seronegative individuals). The present sequence is
CC	human cytomegalovirus lipidated vaccine peptide with covalent HTL and CTL
CC	epitopes
CC	
XX	Sequence 29 AA;
XX	
XX	Query Match 59.3%; Score 83; DB 4; Length 29;
XX	Best Local Similarity 69.2%; Pred. No. 4.2e-06;
XX	Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0
QY	3 QYIKANSKFIGITERGNSLYNTVATL 28
Db	4 QYIKANSKFIGITEAANLVPMVATV 29
XX	
XX	RESULT 11
XX	AAE05437
XX	AAE05437 standard; peptide; 29 AA.
XX	AAE05437;
XX	24-SEP-2001 (first entry)
XX	Human cytomegalovirus (HCMV) vaccine peptide #1.
XX	Human cytomegalovirus; HCMV; infection; vaccine; cytotoxic T lymphocyte;
KW	CTL; CTL precursor; CTLP; Class I antigen;
KM	major histocompatibility complex; MHC; cellular immune response;
KW	CTL activator; CTLP activator; Immunostimulant.
XX	
OS	Human cytomegalovirus.
XX	
XX	Key Location/Qualifiers
FT	Modified-site 1 /note= "Lysine modified with palmitic acid"
FT	Region 4..17 /label= HTL_epitope
FT	Region 18..20 /note= "Amino acid spacer"
FT	Region 21..29 /label= CTL_epitope
XX	
XX	US6251399-B1.
XX	
XX	26-JUN-2001.
XX	
XX	27-MAR-2000; 2000US-00534639.
XX	
XX	12-NOV-1996; 96US-00747488.
PR	14-OCT-1997; 97US-00950064.
PR	10-FEB-1998; 98US-00021298.
PR	11-MAY-1998; 98US-00075257.
XX	
XX	(DIAM/) DIAMOND D J.
PA	(YORK/) YORK J.
XX	
PI	Diamond DJ, York J;
XX	
XX	WPI; 2001-431950/46.
XX	
XX	New peptides that are immunogenic epitopes of the human cytomegalovirus
PT	(CMV), useful for activating cytotoxic T lymphocytes (CTL) or CTL
PT	precursors to elicit an immune response against human CMV by normal or
XX	immunodeficient subjects.
XX	
XX	Example 11; Col 16; 17p; English.
XX	
XX	The invention relates to immunologically active peptides, and functional

PT New immunogenic peptides useful as a vaccination against cytomegalovirus infection, comprises epitopes of human cytomegalovirus which are

Diamond DJ;

XX WPI; 2002-041233/05.
 DR New immunogenic peptides useful as a vaccination against cytomegalovirus
 XX infection, comprises epitopes of human cytomegalovirus which are
 PT recognized by human cytotoxic T-lymphocytes.
 XX
 PS Claim 1; Page 26; 65pp; English.
 XX
 CC The invention relates to immunogenic peptides derived from human
 CC cytomegalovirus (HCMV) which are recognised by human cytotoxic T-
 CC lymphocytes (CTL). The invention also discloses a vaccine against HCMV
 CC comprising one of the peptides, and a pharmaceutically acceptable
 CC carrier, a cellular vaccine against HCMV comprising antigen presenting
 CC cells that have been treated in vitro so that they express the peptide,
 CC and a recombinant viral vector vaccine expressing a gene encoding the
 CC peptide. The vaccine is used to modulate the immune response to human
 CC cytomegalovirus infection. HCMV is implicated in a number of diseases and
 CC conditions including, restenosis after coronary angioplasty,
 CC morbidity/mortality in AIDS (acquired immunodeficiency syndrome)
 CC patients, viral retinitis leading to blindness, encephalitis, enteritis,
 CC mononucleosis, interstitial pneumonia, malignancy (e.g. Kaposi's
 CC sarcoma) and post-bone marrow transplant complications. The present
 CC sequence is an HCMV pp65 based immunogenic epitope lipidated vaccine of
 CC the invention
 XX
 SQ Sequence 29 AA;
 XX
 Query Match 59.3%; Score 83; DB 5; Length 29;
 Best Local Similarity 69.2%; Pred. No. 4.2e-06;
 Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 3 QYIKANSKFIGITGRGRSLYNTVAII 28
 Db 4 QYIKANSKFIGITGRGRSLYNTVAII 29
 XX
 RESULT 14
 AAU11422
 ID AAU11422 standard; peptide; 28 AA.
 XX
 AC AAU11422;
 DT 12-MAR-2002 (first entry)
 XX
 DB Synthetic immunogen peptide 3.
 XX
 KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 XX
 OS Clostridium tetani.
 OS Mammalia.
 OS Syntheetic.
 OS Chimeric.
 XX
 XX Key Location/Qualifiers
 FT 1.15
 FT Peptide /note= "Tetanus toxoid sequence (830-844 aa)"
 FT Peptide 16.19
 FT Peptide /note= "Spacer peptide"
 FT Peptide 20.28
 FT /note= "Gonadotrophin releasing hormone epitope"
 FT Modified-site 28
 FT /note= "Amlidated glycine or glycylamide"
 XX
 XX W0200185763-A2.
 XX
 XX 15-NOV-2001.
 XX
 XX 04-MAY-2001; 2001WO-US014363.
 PF

XX 05-MAY-2000; 2000US-0202328P.
 PR
 XX (APHT-) APHTON CORP.
 PA
 XX Grimes S, Michaeli D, Stevens VC;
 PI
 XX WPI; 2002-049440/06.
 DR
 XX
 XX Novel synthetic immunogen for inducing immune response against
 PT gonadotrophin releasing hormone, comprises fusion peptide having
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 PT or its analog.
 XX
 PS Claim 11; Page 8; 43pp; English.
 XX
 CC The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotrophin releasing hormone (GnRH also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX
 SQ Sequence 28 AA;
 XX
 Query Match 58.6%; Score 82; DB 5; Length 28;
 Best Local Similarity 85.0%; Pred. No. 5.9e-06;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 QYIKANSKFIGITGRGRSLY 22
 Db 1 QYIKANSKFIGITGRGRSLY 20
 XX
 RESULT 15
 ABB79182
 ID ABB79182 standard; peptide; 29 AA.
 XX
 AC ABB79182;
 DT 07-AUG-2002 (first entry)
 XX
 DB Human cytomegalovirus PPI50 related vaccine peptide SEQ ID NO:8.
 DE
 XX Human cytomegalovirus PPI50; HCMV; human leukocyte antigen; HLA; CTL;
 XX cytotoxic T lymphocyte; immunoreactive; epitope; vaccine; infection;
 KW immunogenic; virucide; immune response.
 KW
 XX Human cytomegalovirus.
 OS
 OS W0200234769-A2.
 OS
 XX W0200234769-A2.
 XX
 XX 02-MAY-2002.
 XX
 XX 22-OCT-2001; 2001WO-US032589.
 XX
 XX 20-OCT-2000; 2000US-0241944P.
 XX
 XX (CITY) CITY OF HOPE.
 XX
 XX Diamond DJ;
 PI
 XX
 XX WPI; 2002-471432/50.
 DR
 XX
 XX New immunoreactive peptide cytotoxic T-lymphocytes epitope of human
 PT cytomegalovirus PPI50 useful for creating vaccines against
 PT cytomegalovirus infection.
 PF

line present invention describes a peptide (I) which is an immunogenic epitope recognised by CD8+ class I major histocompatibility complex (MHC) restricted cytotoxic T-lymphocytes (CTL) of patients harbouring latent Cytomegalovirus (HCMV) infection. Also described are: (I) a vaccine (II) against human cytomegalovirus (CMV), which comprises (I); (2) a cellular vaccine (III) against human CMV comprises antigen presenting cells which present (I); (3) a recombinant viral vector (IV) which contains a gene encoding (I); and (4) an immunological reagent (V) which comprises (I). (I) has virulence activity. (II) or (III) are useful for modulating the immune response to HCMV infection, and for vaccinating a mammal in need of it against HCMV. (I) is useful for eliciting a cellular immune response against HCMV, and in immunological methods for creating efficient vaccines against HCMV, and in immunological methods to detect PP150-reactive CTL in a patient or a sample from the patient. The present sequence represents a HCMV PP150 related vaccine peptide, which is given in the exemplification of the present invention

```

QY      1 GRÖYIKANSKFIGITE 16
         |||||
Db      14 GRÖYIKANSKFIGITE 29

```

Search completed: July 20, 2004, 06:30:26
Job time : 43.7705 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:10:18 ; Search time 9.63934 Seconds

(Without alignments)
279,413 Million cell updates/sec

Title: US-09-673-166B-274

Perfect score: 144

Sequence: 1 GRQYIKANSKFIGITERGRILKEPVHGV 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	48.6	1315	1 BRLTN	tentoxilysin (EC 3
2	54	37.5	559	2 B47175	reverse transcript
3	54	37.5	1002	2 S54378	pol polypeptide
4	54	37.5	1003	1 GNWMLV	HIV-1 retropepsin
5	54	37.5	1003	1 B44001	HIV-1 retropepsin
6	54	37.5	1003	2 T09440	pol polypeptide
7	54	37.5	1012	1 GNWMLV	HIV-1 retropepsin
8	54	37.5	1015	1 GNWMLV	HIV-1 retropepsin
9	53	36.8	1002	1 GNWMLV	HIV-1 retropepsin
10	52	36.1	902	2 T01668	HIV-1 retropepsin
11	51	35.4	559	2 A47175	reverse transcript
12	50	34.7	243	2 P96595	unknown protein, 2
13	49	34.0	278	2 G82074	mag2 protein VCC45
14	49	34.0	323	2 P95877	probable ABC trans
15	48	33.3	83	2 B84174	hypothetical prote
16	48	33.3	492	2 E86829	hypothetical prote
17	48	33.3	697	1 TVVPTL	large ABC transpo
18	47	32.6	152	2 A37815	globin (domain E7)
19	47	32.6	352	2 H1284	probable o-sialogl
20	47	32.6	674	2 T20571	hypothetical prote
21	47	32.6	721	2 C84073	DNA topoisomerase
22	46.5	32.3	227	2 AH1667	phosphoribosylform
23	46.5	32.3	287	2 D97164	ATPases involved i
24	46.5	32.3	564	2 T28949	hypothetical prote
25	46.5	32.3	899	2 T42976	hypothetical prote
26	46	31.9	447	2 T43043	probable acetyl-Co
27	46	31.9	447	2 H97146	sldephore/Surfac
28	46	31.9	448	2 B64503	hypothetical prote
29	46	31.9	493	2 H69689	ribose ABC transp

30	46	31.9	511	2 E84685	probable seed stor
31	46	31.9	569	2 B82559	30S ribosomal prot
32	46	31.9	728	2 I65413	sodium-dependent n
33	46	31.9	730	2 I52632	sodium-dependent n
34	46	31.9	912	2 S33980	pol polypeptide
35	46	31.9	922	2 AG1827	maltooligosyltreha
36	46	31.9	1003	1 GNWMA2	HIV-1 retropepsin
37	45.5	31.6	455	2 T10406	protein ME53 - Org
38	45.5	31.6	899	2 G36812	hypothetical prote
39	45	31.2	247	2 F47188	MHC class II histo
40	45	31.2	286	2 A95113	cellulite resistan
41	45	31.2	288	2 H97981	telurite resistan
42	45	31.2	424	2 G84317	hypothetical prote
43	45	31.2	462	2 G96506	hypothetical prote
44	45	31.2	527	2 T37055	probable oxidoredu
45	45	31.2	689	2 E89898	hypothetical prote

ALIGNMENTS

RESULT 1
BRLTN
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N:Alternate names: tetanus neurotoxin
C:Species: Clostridium tetani
C>Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 03-Jun-2002
C/Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364
R:Eisel, U.; Jarasch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, M.
EMBO J. 5, 2495-2502, 1986
A>Title: Tetanus toxin: Primary structure, expression in E. coli, and homology with botu
A:Reference number: A25689; MUID:87053814; PMID:3536478
A:Accession: A25689
A:Molecule type: DNA
A:Residues: 1-1315 <ERS>
A:Cross-references: GB:X04436; NID:940769; PIDN:CAA28033.1; PID:940770
R:Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A>Title: The complete nucleotide sequence of tetanus toxin.
A:Reference number: A25757; MUID:87040747; PMID:3774547
A:Accession: A25757
A:Molecule type: DNA
A:Residues: 1-1315 <FAI>
A:Cross-references: GB:X06214; NID:940773; PIDN:CAA29564.1; PID:940774
R:Experimental source: strain CN3911
R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A>Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in E
A:Reference number: A25194; MUID:86085672; PMID:3510187
A:Accession: A25194
A:Molecule type: DNA
A:Residues: 743-1315 <FA2>
A:Cross-references: GB:M12739; NID:9144920; PIDN:AAA23282.1; PID:9144921
A:Accession: B25194
A:Molecule type: protein
A:Residues: 865-894 <FA3>
R:Matsuda, M.; Iel, D.L.; Sugimoto, N.; Ozutsuni, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A>Title: Isolation, purification, and characterization of fragment B, the NH-2-terminal
A:Reference number: A60759; MUID:90035436; PMID:2478476
A:Accession: A60759
A:Molecule type: protein
A:Residues: 461-475 <MAT>
R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A>Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
A:Reference number: JS0098; MUID:89093918; PMID:2463305
A:Contents: annotation; epitope region
R:Schiano, G.; Benfenati, F.; Pouliat, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta, B.R.
Nature 359, 832-835, 1992
A>Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic
A:Reference number: S27125; MUID:93063293; PMID:1331807
A:Contents: annotation

R:de Filipiis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
 Bur. J. Biochem. 229, 61-69, 1995
 A:Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
 A:Reference number: S69348; MUID:95262688; PMID:7744050
 A:Accession: S69348
 A:Molecule type: Protein
 A:Residues: 2-31 <DEF>
 C:Comment: The source of this protein was an extrachromosomal plasmid.
 C:Comment: The precursor is cleaved by endogenous proteinase activity to form light (fr
 dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)
 C:Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to gangli
 C:Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized
 presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy
 C:Function:
 A:Description: blocks neuroexocytosis via hydrolysis of a Glu-Phe peptide bond in synapt
 C:Superfamily: tetanus toxin
 C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
 F:2-457/Product: tetroxylisin light chain (fragment A) #status predicted <TRL>
 F:461-1315/Product: tetroxylisin heavy chain (fragment B.C) #status experimental <TTH>
 F:461-864/Dmain: channel forming (fragment B) #status predicted <TRX>
 F:865-1315/Dmain: ganglioside binding (fragment C) #status predicted <TXC>
 F:233-237/Binding site: zinc (His) #status predicted
 F:234/Active site: Glu #status predicted

Query Match 48.6%; Score 70; DB 1; Length 1315;
 Best Local Similarity 100.0%; Pred. No. 0.067;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QYKANSKFIIGITE 16
 |||||
 Db 830 QYKANSKFIIGITE 843

RESULT 2

B47175
 Reverse transcriptase, AZT-sensitive variant - human immunodeficiency virus type 1
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 13-Sep-1996
 C:Accession: B47175
 R:Moht, H.; Singh, M.K.; Ching, W.T.; Ho, D.D.
 Proc. Natl. Acad. Sci. U.S.A. 90, 25-29, 1993
 A:Title: Quantitation of zidovudine-resistant human immunodeficiency virus type 1 in the
 A:Reference number: A47175; MUID:93126353; PMID:7678340
 A:Accession: B47175
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: DNA
 A:Note: Sequence extracted from NCBI backbone (NCBIF:122099)
 C:Superfamily: pol polypeptide

Query Match 37.5%; Score 54; DB 2; Length 559;
 Best Local Similarity 45.2%; Pred. No. 6.1;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIIGITERG-----RIKKEPVHGV 28
 |||||
 Db 287 KALTEVIPLEAELELAENREILKEPVHGV 317

RESULT 3

S54378
 pol polypeptide - human immunodeficiency virus type 1 (fragment)
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 26-Aug-1999
 C:Accession: S54378
 R:Theodore, T.; Buckler-White, A.J.
 submitted to the EMBL Data Library, July 1989
 A:Reference number: S54377
 A:Accession: S54378
 A:Status: preliminary
 A:Molecule type: genomic RNA
 A:Residues: 1-1002 <THE>
 A:Cross-references: EMBL:M2639; NID:g329377; PIDN:AAA45366.1; PID:g329381

C:Superfamily: pol polypeptide
 C:Keywords: polypeptide

Query Match 37.5%; Score 54; DB 2; Length 1002;
 Best Local Similarity 45.2%; Pred. No. 11;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIIGITERG-----RIKKEPVHGV 28
 |||||
 Db 441 KALTEVIPLEAELELAENREILKEPVHGV 471

RESULT 4

GNVWIV
 HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate LAV-1a)
 N:Contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA polym
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 03-Jun-2002
 C:Accession: A03966
 R:Main-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.
 Cell 40, 9-17, 1985
 A:Title: Nucleotide sequence of the AIDS virus, LAV.
 A:Reference number: A90865; MUID:8509333; PMID:2981635
 A:Accession: A03966
 A:Molecule type: DNA
 A:Residues: 1-1003 <MAI>
 C:Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
 C:Genetics:
 A:Gene: pol
 C:Superfamily: pol polypeptide
 C:Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleot
 F:57-155/Product: retropepsin #status predicted <RTP>
 F:81/Active site: Asp (shared with dimeric partner) #status experimental

Query Match 37.5%; Score 54; DB 1; Length 1003;
 Best Local Similarity 45.2%; Pred. No. 11;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIIGITERG-----RIKKEPVHGV 28
 |||||
 Db 442 KALTEVIPLEAELELAENREILKEPVHGV 472

RESULT 5

B44001
 HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (strain YU-2)
 N:Contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA polym
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
 C:Accession: B44001
 R:Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.
 J. Virol. 66, 6587-6600, 1992
 A:Title: Complete nucleotide sequence, genome organization, and biological properties of
 A:Reference number: A44001; MUID:93021387; PMID:1404605
 A:Accession: B44001
 A:Molecule type: DNA
 A:Residues: 1-1003 <LIV>
 A:Cross-references: GB:M93258
 C:Comment: This protein is synthesized as a gag-pol polypeptide.
 C:Comment: Specific enzymatic cleavages may yield mature proteins including proteinase, i
 C:Genetics:
 A:Gene: pol
 C:Superfamily: pol polypeptide
 C:Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleot
 F:57-155/Product: retropepsin #status predicted <RTP>
 F:81/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 37.5%; Score 54; DB 1; Length 1003;
 Best Local Similarity 45.2%; Pred. No. 11;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

```

Qy      6 KANSKFIGITRG-----RIIKEPVHGV 28
      ||::|||
      A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.
      A:Reference number: A93553; MUID:85111123; PMID:2578615
      A:Accession: A03965
      A:Molecule type: DNA
      A:Residues: 1-1015 <RAT>
      A:Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:9326383; PIDN:AAA4445
      A:Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
      C:Gene: pol
      C:Superfamily: pol polyprotein

RESULT 6
T09440
pol polyprotein - human immunodeficiency virus type 1 (strain JRP1) (fragment)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-Jan-2000
C:Accession: T09440
R:Yang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,
Submitted to the EMBL Data Library, July 1996
A:Reference number: Z16673
A:Accession: T09440
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1003 <PAN>
A:Cross-references: EMBL:U63632; NID:91465777; PID:91465779
C:Gene: pol
C:Superfamily: pol polyprotein

Query Match      37.5%; Score 54; DB 2; Length 1003;
Best Local Similarity 45.2%; Pred. No. 11;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy      6 KANSKFIGITRG-----RIIKEPVHGV 28
      ||::|||
      Db      442 KALTEVPLTEAELELAENREILKEPVHGV 472

RESULT 7
GNVWVL
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate IV)
N:Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 03-Jun-2002
C:Accession: A03967
R:Musling, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.
Nucleotide sequence of HIV-1 RNA: a highly cytopathic strain of the human immuno
A:Title: Nucleotide acid structure and expression of the human AIDS/lymphadenopathy retrovi
A:Reference number: A93355; MUID:85111157; PMID:2982104
A:Accession: A03967
A:Molecule type: DNA
A:Residues: 1-1012 <MUE>
A:Cross-references: GB:M27323; NID:9328154; PIDN:AAA44869.1; PID:9328158
C:Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
C:Gene: pol
C:Superfamily: pol polyprotein
C:Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleo
F:66-164/Product: retropepsin #status predicted <RTP>
F:90/Active site: Asp (shared with dimeric partner) #status experimental

Query Match      37.5%; Score 54; DB 1; Length 1012;
Best Local Similarity 45.2%; Pred. No. 12;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy      6 KANSKFIGITRG-----RIIKEPVHGV 28
      ||::|||
      Db      451 KALTEVPLTEAELELAENREILKEPVHGV 481

RESULT 8
GNVWHL
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate HTIV-III
N:Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 03-Jun-2002
C:Accession: A03965
R:Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Dora
nberger, U.A.; Papas, T.S.; Ghayab, U.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.

```

```

      Nature 313, 277-284, 1985
      A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.
      A:Reference number: A93553; MUID:85111123; PMID:2578615
      A:Accession: A03965
      A:Molecule type: DNA
      A:Residues: 1-1015 <RAT>
      A:Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:9326383; PIDN:AAA4445
      A:Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
      C:Gene: pol
      C:Superfamily: pol polyprotein
      C:Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucle
      F:69-167/Product: retropepsin #status predicted <RTP>
      F:93/Active site: Asp (shared with dimeric partner) #status experimental

Query Match      37.5%; Score 54; DB 1; Length 1015;
Best Local Similarity 45.2%; Pred. No. 12;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy      6 KANSKFIGITRG-----RIIKEPVHGV 28
      ||::|||
      Db      454 KALTEVPLTEAELELAENREILKEPVHGV 484

RESULT 9
GNLJND
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate NDK)
N:Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 03-Jun-2002
C:Accession: J00067
R:Spire, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe, A.;
Gene 81, 275-286, 1989
A:Title: Nucleotide sequence of HIV-1 NDK: a highly cytopathic strain of the human immuno
A:Reference number: J00065; MUID:90034200; PMID:2806917
A:Accession: J00067
A:Molecule type: DNA
A:Residues: 1-1002 <SPI>
A:Cross-references: GB:M27323; NID:9328154; PIDN:AAA44869.1; PID:9328158
C:Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
C:Gene: pol
C:Superfamily: pol polyprotein
C:Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleo
F:56-154/Product: retropepsin #status predicted <RTP>
F:80/Active site: Asp (shared with dimeric partner) #status predicted

Query Match      36.8%; Score 53; DB 1; Length 1002;
Best Local Similarity 41.9%; Pred. No. 16;
Matches 13; Conservative 4; Mismatches 6; Indels 8; Gaps 1;

Qy      6 KANSKFIGITRG-----RIIKEPVHGV 28
      ||::|||
      Db      441 KALTEVPLTEAELELAENREILKEPVHGV 471

RESULT 10
T01668
pol polyprotein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T01668
R:Alizon, M.; Main-Hobson, S.; Gluckman, J.C.; Sonigo, P.
Cell 46, 63-74, 1986
A:Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two isolat
A:Reference number: Z14389; MUID:86245056; PMID:2424612
A:Accession: T01668
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-902 <ALI>
A:Cross-references: EMBL:K03456; NID:960228; PIDN:CAA28012.1; PID:960230
C:Superfamily: pol polyprotein

```

Query Match 36.1%; Score 52; DB 2; Length 902;
 Best Local Similarity 66.7%; Pred. No. 20;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

RESULT 11

A47175
 reverse transcriptase, AZT-resistant variant - human immunodeficiency virus type 1
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 13-Sep-1996
 C:Accession: A47175
 R:Mohtai, H.; Singh, M.K.; Ching, W.T.; Ho, D.D.
 Proc. Natl. Acad. Sci. U.S.A. 90, 25-29, 1993
 A:Title: Quantitation of zidovudine-resistant human immunodeficiency virus type 1 in the
 A:Reference number: A47175; MUID:93126353; PMID:7678340
 A:Accession: A47175
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: nucleic acid
 A:Residues: 1-559 <MOH>
 A:Note: sequence extracted from NCBI backbone (NCBIP:122100)
 C:Superfamily: pol polypeptide

Query Match 35.4%; Score 51; DB 2; Length 559;
 Best Local Similarity 41.9%; Pred. No. 17;
 Matches 13; Conservative 4; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKPIGTERG-----RIKEPVHG 28
 DB 287 KALTEVQLTEFAELELAENREILREPVG 317

RESULT 12

P96595
 unknown protein, 25817-24837 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: P96595
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conu, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, U.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
 Kerr, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: P96595
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-243 <STO>
 A:Cross-references: GB:AE005173; NID:G10645423; PIDN:AA621540.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F7A10.13
 A:Map position: 1

Query Match 34.7%; Score 50; DB 2; Length 243;
 Best Local Similarity 50.0%; Pred. No. 9.9;
 Matches 12; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 2 RQYIKANSKFIIGTERGRILKEPV 25
 DB 94 RYTTTRGS--WGIDSGRIKEPV 115

RESULT 13

GB2074

mazG protein VC2450 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 18-Aug-2003
 C:Accession: G82074
 R:Heidelberger, J.F.; Bisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.
 L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: G82074
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-278 <HRT>
 A:Cross-references: GB:AE004315; GB:AE003852; NID:G9657024; PIDN:AAF95592.1; GSPDB:GN0012
 A:Experimental source: serogroup O1, strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC2450
 A:Map position: 1
 C:Superfamily: nucleoside triphosphate pyrophosphohydrolase, MazG type; beta-lactamase re

Query Match 34.0%; Score 49; DB 2; Length 278;
 Best Local Similarity 36.0%; Pred. No. 16;
 Matches 9; Conservative 7; Mismatches 5; Indels 4; Gaps 1;

QY 6 KANSKPI-----GITERGRILKEPVH 26
 DB 228 KANAKFVRFRFGVEDEKAKKKPKH 252

RESULT 14

E95877
 probable ABC transporter periplasmic substrate-binding protein SMD20295 [imported] - Sinc
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 01-Mar-2002
 C:Accession: E95877
 R:Finan, T.M.; Weidner, S.; Wong, K.; Bhurmaster, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: E95877
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-323 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC48685.1; PID:G15140157; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymB
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Aboia, P.; Ampe, F.; Batloy-Hubler,
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebnult, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMD20295
 A:Genome: Plasmid
 C:Superfamily: conserved hypothetical protein H11028

Query Match 34.0%; Score 49; DB 2; Length 323;
 Best Local Similarity 40.7%; Pred. No. 19;
 Matches 11; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 GROVTKANSKFIIGTERGRILKEPVHG 27
 DB 30 GMOGTAGDPQPEGVTEARITIKESGG 56

RESULT 15

B84174
 hypothetical protein Vng0138h [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C;Accession: B84174
 R;Ng, W.V.; Kennedy, S.P.; Mahlitz, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
 ; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; L
 A;Title: Genome sequence of Halobacterium species NRC-1.
 A;Reference number: A84160; MUID:20504483; PMID:11016950
 A;Accession: B84174
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-83 <STO>
 A;Cross-references: GB:AE004437; NID:G10579784; PIDN:AA618758.1; GSPDB:GNC0138
 C;Genetics:
 A;Gene: VNG0138H

Query Match 33.3%; Score 48; DB 2; Length 83;
 Best Local Similarity 50.0%; Pred. No. 6.2;
 Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 2 RQYXANSKFFIGTERGR 19
 | : : : : : | | | | |
 Db 35 RRAVEASAKVTGAERGR 52

Search completed: July 20, 2004, 06:34:10
 Job time : 11.6393 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:17 ; Search time 5.5082 Seconds

(without alignments)
264.690 Million cell updates/sec

Title: US-09-673-166B-274

Sequence: 144
1 GROYIKANSKFGITGRGRIKPEPHGV 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	46.6	1314	1	TEXX_CLOTE
2	54	37.5	1002	1	P04958 clostridium
3	54	37.5	1002	1	P04589 human immun
4	54	37.5	1003	1	P12499 human immun
5	54	37.5	1003	1	P04585 human immun
6	54	37.5	1003	1	P20892 human immun
7	54	37.5	1003	1	P35963 human immun
8	54	37.5	1006	1	P03366 human immun
9	54	37.5	1015	1	P03367 human immun
10	54	37.5	1015	1	P03368 human immun
11	54	37.5	1015	1	P14802 human immun
12	53	36.8	1002	1	P14997 bacillus su
13	53	36.8	1003	1	P04588 human immun
14	52	36.1	1002	1	P05959 human immun
15	52	36.1	1002	1	P20875 human immun
16	52	36.1	1007	1	P08740 pyrobaculum
17	49	34.0	1007	1	P24740 human immun
18	49	34.0	1002	1	P03367 human immun
19	48	33.3	697	1	TALA_POVLY
20	47	32.6	152	1	GLB7_ARTSX
21	47	32.6	352	1	GCP_TREPA
22	46.5	32.3	227	1	PUR0_LISTN
23	46	31.9	448	1	Y628_MERTA
24	46	31.9	493	1	RBSA_BACV
25	46	31.9	729	1	NTT7_BOVIN
26	46	31.9	729	1	NTT7_RAT
27	46	31.9	730	1	NTT7_HUMAN
28	46	31.9	1003	1	POL_HV1A2
29	45.5	31.6	455	1	ME53_NPVOP
30	45.5	31.6	899	1	V120_HSVSA
31	45	31.2	207	1	PHI_BUCBP
32	45	31.2	384	1	LE12_THETN
33	44.5	30.9	288	1	ALF_MYCGE

34	44.5	30.9	585	1	SP2E_BACME	P49600 bacillus me
35	44.5	30.9	1682	1	MSPI_PLAFC	P19588 plasmodium
36	44.5	30.9	1701	1	MSPI_PLAFC	P13819 plasmodium
37	44.5	30.9	1701	1	MSPI_PLAFC	P08569 plasmodium
38	44.5	30.9	1726	1	MSPI_PLAFC	P04934 plasmodium
39	44.5	30.9	1726	1	MSPI_PLAFC	P04935 plasmodium
40	44	30.6	66	1	VG84_BPMLS	005301 mycobacteri
41	44	30.6	333	1	ILVC_MERTH	027491 methanobact
42	44	30.6	354	1	RECG_ACTFA	P34716 acidiphiliu
43	44	30.6	364	1	ARGD_PYRAB	Q9v114 pyrococcus
44	44	30.6	660	1	TPLE_XLFA	Q9phk2 xylella fas
45	43.5	30.2	106	1	Y094_HAEIN	P43939 haemophilus

ALIGNMENTS

RESULT 1
ID TEXX_CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin) [Contains:
DE tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy
DE chain (Tetanus toxin chain H)].
GN TEXX OR CYP60.
OS Clostridium tetani.
OG Plasmid p88, and Plasmid 75 Kbp.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=87053814; PubMed=3536478;
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudek M., Habermann E., Niemann H.,
RT "Tetanus toxin: primary structure, expression in E. coli, and
RL homology with botulinum toxins.";
RN EMBO J. 5:2495-2502(1986).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CN3911; PLASMID=75 Kbp;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.,
RT "The complete nucleotide sequence of tetanus toxin.";
RL Nucleic Acids Res. 14:7809-7812(1986).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88; PLASMID=PE88;
RX MEDLINE=22457253; PubMed=12552129;
RA Brieglebmann H., Baumeister S., Fricke W.F., Wierzer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.,
RT "The genome sequence of Clostridium tetani, the causative agent of
RL tetanus disease.";
RN Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
[4]
RP SEQUENCE OF 742-1314 FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=6085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.,
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT fragment C in Escherichia coli.";
RN J. Bacteriol. 165:21-27(1986).
[5]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90201034; PubMed=2108021;
RA Kriegstein K., Henschen A., Weller U., Habermann E.,
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
in tetanus toxin.";

RL Eur. J. Biochem. 188:39-45(1990).
 RN
 RP PARTIAL SEQUENCE.
 RX MEDLINE=92037649; PubMed=1935979;
 RA Krieglstein K.G., Henschel A.H., Weller U., Habermann E.;
 RT "Limited proteolysis of tetanus toxin. Relation to activity and
 RT identification of cleavage sites.";
 RL Eur. J. Biochem. 202:41-51(1991).
 RN
 RP IDENTIFICATION AS ZINC-PROTEASE.
 RX MEDLINE=93010948; PubMed=1396558;
 RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
 RA Montecucco C.;
 RT "Tetanus toxin is a zinc protein and its inhibition of
 RT neurotransmitter release and protease activity depend on zinc.";
 RL EMBO J. 11:3577-3583(1992).
 RN
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331807;
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
 RT by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835(1992).
 RN
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
 RX MEDLINE=97475217; PubMed=9344741;
 RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
 RA Sax M.;
 RT "Structure of the receptor binding fragment HC of tetanus
 RT neurotoxin.";
 RL Nat. Struct. Biol. 4:788-792(1997).
 CC
 CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN|-PHE-77
 CC BOND OF SYNAPTOSOMAL-2.
 CC
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln|-Phe-77 bond in
 CC synaptobrevin 2.
 CC
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (by similarity).
 CC
 CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
 CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
 CC AND ARE NON-TOXIC AFTER SEPARATION.
 CC
 CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
 CC GANGLIOSIDE RECEPTORS.
 CC
 CC -1- SIMILARITY: Belongs to peptidase family M27.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X04436; CAA28033.1; -
 DR EMBL: X06214; CAA29564.1; -
 DR EMBL: AF528097; AAC37454.1; -
 DR EMBL: M12739; AAA23282.1; -
 DR PIR: A25689; BTUTIN.
 DR PDB: 1AF9; 29-APR-98.
 DR PDB: 1ABD; 14-OCT-98.
 DR PDB: 1DOH; 27-MAR-00.
 DR PDB: 1DFO; 24-MAR-00.
 DR PDB: 1DIW; 24-MAR-00.
 DR PDB: 1DLI; 24-MAR-00.
 DR PDB: 1FV3; 05-SEP-01.
 DR MEROPS: M27.001; -
 DR InterPro: IPR008985; Cona like lec_g1.
 DR InterPro: IPR002160; Kunitz legume.
 DR InterPro: IPR006025; Pept_M_Zn_BS.

DR InterPro: IPR000395; Peptidase M27.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOKLYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KM Neurotoxin; Transmembrane; Hydrolyase; Metalloprotease; Zinc; Plasmid;
 KW 3D-structure; Complete proteome.
 FT INIT_MET 0
 FT CHAIN 1
 FT CHAIN 456
 FT METAL 1314
 FT METAL 232
 FT ACT_SITE 233
 FT METAL 236
 FT TRANSMEM 226
 FT TRANSMEM 669
 FT DISULFID 438
 FT DISULFID 1076
 FT HELIX 876
 FT TURN 882
 FT TURN 883
 FT STRAND 884
 FT TURN 884
 FT TURN 884
 FT STRAND 884
 FT STRAND 892
 FT STRAND 897
 FT TURN 904
 FT TURN 909
 FT STRAND 912
 FT STRAND 915
 FT STRAND 920
 FT TURN 928
 FT TURN 929
 FT STRAND 932
 FT STRAND 935
 FT HELIX 940
 FT TURN 941
 FT STRAND 949
 FT STRAND 956
 FT HELIX 962
 FT TURN 968
 FT TURN 969
 FT TURN 972
 FT STRAND 972
 FT STRAND 980
 FT HELIX 983
 FT STRAND 987
 FT TURN 996
 FT STRAND 998
 FT TURN 1004
 FT TURN 1006
 FT STRAND 1010
 FT STRAND 1016
 FT TURN 1020
 FT TURN 1021
 FT STRAND 1031
 FT TURN 1039
 FT STRAND 1042
 FT TURN 1048
 FT TURN 1050
 FT STRAND 1058
 FT TURN 1068
 FT STRAND 1079
 FT TURN 1082
 FT STRAND 1091
 FT HELIX 1097
 FT TURN 1106
 FT STRAND 1112
 FT STRAND 1114
 FT TURN 1116
 FT STRAND 1120
 FT STRAND 1122
 FT TURN 1123
 FT TURN 1127
 FT HELIX 1132
 FT TURN 1135
 FT STRAND 1137
 FT TURN 1144
 FT TURN 1144
 FT STRAND 1148
 FT STRAND 1152
 FT STRAND 1155
 FT TURN 1159
 FT STRAND 1162
 FT STRAND 1163
 FT STRAND 1173
 FT TURN 1178
 FT TURN 1184
 FT STRAND 1185
 FT STRAND 1188
 FT STRAND 1190

Query Match 48.6%; Score 70; DB 1; Length 1314;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 QYIKANSKFIGITE 16
|||
Db 829 QYIKANSKFIGITE 842

RESULT 2
POL_HV1EL STANDARD; PRT; 1002 AA.
ID POL_HV1EL
AC P04589; 077906;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1988 (Rel. 09, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (H1 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86245056; PubMed=2424612;
RA Alison M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients."
RL Cell 46:63-74(1986).
CC -1- CATALYTIC ACTIVITY: Specific for a p1 residue that is hydrophobic,
and p1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ [DNA] (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC

DR EMBL; A07108; CA00612.1; -;
DR EMBL; K03454; AAA44325.1; ALT_INIT.
DR HSSP; P03366; IHMV.
DR HIV; K03454; POLSELI.
DR MEROPS; A02.001; -;
DR InterPro; IPR001969; Aspprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rye.
DR InterPro; IPR000477; RYase.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
KW AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferase; RNA-directed DNA polymerase.
FT CHAIN 56 154 PROTEASE.

FT ACT SITE 80 80 BY SIMILARITY.
SQ SEQUENCE 1002 AA; 114002 MW; 5700903B6893B54 CRC64;
Query Match 37.5%; Score 54; DB 1; Length 1002;
Best Local Similarity 45.2%; Pred. No. 3.3;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
QY 6 KANSKFIGTERG-----RIKEPVHGV 28
|||
Db 441 KALTEVIFLTERAELELAENRILKEPVHGV 471

RESULT 3
POL_HV1Z2 STANDARD; PRT; 1002 AA.
ID POL_HV1Z2
AC P12439;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (Z2/CDC-Z34 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11683;
RN [1]
RP SEQUENCE FROM N.A.
RA Theodore T., Buckler-White A.;
RT Submitted (NOV-1988) to the HIV data bank.
RL -1- CATALYTIC ACTIVITY: Specific for a p1 residue that is hydrophobic,
and p1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ [DNA] (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC

DR EMBL; M22639; AAA45366.1; -;
DR PIR; S54378; S54378.
DR PDB; 1E28; 12-SEP-00.
DR PDB; 1HXW; 04-FEB-98.
DR PDB; 1PRO; 17-AUG-96.
DR PDB; 1V1U; 13-JAN-99.
DR PDB; 1V1K; 13-JAN-99.
DR HIV; M22639; POL\$Z2Z6.
DR MEROPS; A02.001; -;
DR InterPro; IPR001969; Aspprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rye.
DR InterPro; IPR000477; RYase.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.

KW AIDS; Polypeptide; Hydrolase; Aspartyl protease; Endonuclease;
 KM Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
 FT ACT SITE 56 154 BY SIMILARITY.
 FT ACT SITE 80 80 BY SIMILARITY.
 SQ SEQUENCE 1002 AA; 113724 MW; CB4AAC9AB4742315 CRC64;
 Query Match 37.5%; Score 54; DB 1; Length 1002;
 Best Local Similarity 45.2%; Pred. No. 3.3;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKRGIGTERG-----RIIKRPYHGV 28
 Db 441 KALTEVIFPLEEAEHLEANNRIKKEPVHGV 471
 RESULT 4
 ID POL_HV1H2 STANDARD; PRT; 1003 AA.
 AC P04585; Q09777; Q9W0C5;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE POL. polypeptide [contains: Protease (Retropesin) (EC 3.4.23.16);
 DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN POL.
 OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OC NCBI_TaxID=11706;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=872929196; PubMed=3040055;
 RA Ogata N., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
 RA Gallo R.C., Wong-Staal F.;
 RT "Complete nucleotide sequences of functional clones of the AIDS
 RT virus.";
 RT AIDS Res. Hum. Retroviruses 3:57-69(1987).
 RN [2]
 RP REVISIONS.
 RX Ogata N., Alter H.J., Miller R.H., Purcell R.H.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chappey C.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 57-155.
 RX MEDLINE=99043699; PubMed=9827997;
 RA Kervinen J., Lubkowski J., Zdanov A., Bhatt D., Dunn B.M., Hu K.Y.,
 RA Powell D.J., Kay J., Wlodawer A., Gustchina A.;
 RT "Toward a universal inhibitor of retroviral proteases: comparative
 RT analysis of the interactions of LP-130 complexed with proteases from
 RT HIV-1, FIV, and EIAV.";
 RL Protein Sci. 7:2314-2323(1998).
 RN [5]
 RP STRUCTURE BY NMR OF 57-155.
 RX MEDLINE=97022126; PubMed=866486;
 RA Yamazaki T., Hinck A.P., Wang Y.-X., Nicholson L.K., Torchia D.A.,
 RA Wingfield P., Stahl S.J., Kaufman J.D., Chang C.-H., Donatelli P.U.,
 RA Lam P.Y.S.;
 RT "Three-dimensional solution structure of the HIV-1 protease complexed
 RT with DMP323, a novel cyclic urea-type inhibitor, determined by
 RT nuclear magnetic resonance spectroscopy.";
 RL Protein Sci. 5:495-506(1996).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 156-595.
 RX MEDLINE=96097398; PubMed=8535785;
 RA Ren J., Esnouf R.M., Hopkins A.L., Ross C., Jones E.Y., Stammers D.K.,
 RA Stuart D.I.;
 RT "The structure of HIV-1 reverse transcriptase complexed with
 RT 9-chloro-TIBO: lessons for inhibitor design.";
 RL Structure 3:915-926(1995).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.55 ANGSTROMS) OF 156-595.

RX MEDLINE=96208551; PubMed=8648598;
 RA Hopkins A.L., Ren J., Esnouf R.M., Wilcox B.E., Jones E.Y., Ross C.,
 RA Miyasaka T., Walker R.T., Tanaka H., Stammers D.K., Stuart D.I.;
 RT "Complexes of HIV-1 reverse transcriptase with inhibitors of the HEPT
 RT series reveal conformational changes relevant to the design of potent
 RT non-nucleoside inhibitors.";
 RL J. Med. Chem. 39:1589-1600(1996).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 157-694.
 RX MEDLINE=97268683; PubMed=9108091;
 RA Esnouf R.M., Ren J., Hopkins A.L., Ross C.K., Jones E.Y.,
 RA Stammers D.K., Stuart D.I.;
 RT "Unique features in the structure of the complex between HIV-1
 RT reverse transcriptase and the bis(heteroaryl)piperazine (BHP)
 RT U-90152 explain resistance mutations for this nonnucleoside
 RT inhibitor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:3984-3989(1997).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 157-694.
 RX MEDLINE=98356189; PubMed=9689112;
 RA Ren J., Esnouf R.M., Hopkins A.L., Jones E.Y., Kirby I., Keeling J.,
 RA Ross C.K., Larder B.A., Stuart D.I., Stammers D.K.;
 RT "3'-Azido-3'-deoxythymidine drug resistance mutations in HIV-1
 RT reverse transcriptase can induce long range conformational changes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:9518-9523(1998).
 CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
 CC and P1 variable, but often pro.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphomonoester.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + [DNA] (N).
 CC -1- Ptm: Cleavage sites that yield the mature proteins remain to be
 CC determined.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 CC KNOWN AS THE RETROPEPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; K03455; AAB50259.1; ALT_INIT.
 DR EMBL; AF033819; AAC82598.2; -.
 DR PDB; 1REV; 14-OCT-96.
 DR PDB; 1RT1; 21-APR-97.
 DR PDB; 1RT2; 21-APR-97.
 DR PDB; 1RT3; 16-FEB-99.
 DR PDB; 1RT4; 03-APR-96.
 DR PDB; 1RT1; 03-APR-96.
 DR PDB; 1RT0; 03-APR-96.
 DR PDB; 1BVE; 17-AUG-96.
 DR PDB; 1BVG; 17-AUG-96.
 DR PDB; 1KLM; 18-MAR-98.
 DR PDB; 1A30; 29-APR-98.
 DR PDB; 1OD1; 16-FEB-99.
 DR PDB; 1BV7; 14-JAN-00.
 DR PDB; 1BV9; 12-JAN-00.
 DR PDB; 1BWA; 12-JAN-00.
 DR PDB; 1BWB; 12-JAN-00.
 DR PDB; 1COT; 19-JUL-00.
 DR PDB; 1COT; 19-JUL-00.
 DR PDB; 1COT; 19-JUL-00.
 DR PDB; 1C1B; 21-JUL-00.
 DR PDB; 1C1C; 06-SEP-00.
 DR PDB; 1C1D; 06-SEP-00.
 DR PDB; 1DMP; 12-NOV-97.
 DR PDB; 1DT0; 20-MAR-00.
 DR PDB; 1DTT; 02-APR-00.
 DR PDB; 1EP4; 27-SEP-00.
 DR PDB; 1EX4; 26-JUL-00.
 DR PDB; 1EXQ; 03-NOV-00.
 DR PDB; 1FK9; 28-FEB-01.

DR PDB; 1FKO; 28-FEB-01.
 DR PDB; 1FKP; 03-NOV-00.
 DR PDB; 1HVA; 06-APR-99.
 DR PDB; 1HVR; 15-MAY-95.
 DR PDB; 1HWR; 23-MAR-99.
 DR PDB; 1HXB; 12-MAR-97.
 DR PDB; 1JXH; 03-OCT-01.
 DR PDB; 1JLA; 03-OCT-01.
 DR PDB; 1JLB; 03-OCT-01.
 DR PDB; 1JLC; 03-OCT-01.
 DR PDB; 1JLF; 03-OCT-01.
 DR PDB; 1JLG; 03-OCT-01.
 DR PDB; 1JLQ; 22-AUG-01.
 DR PDB; 1LM0; 30-OCT-02.
 DR PDB; 1LM2; 30-OCT-02.
 DR PDB; 1LMC; 30-OCT-02.
 DR PDB; 1LWE; 30-OCT-02.
 DR PDB; 1LWF; 30-OCT-02.
 DR PDB; 1MER; 15-APR-98.
 DR PDB; 1MES; 15-APR-98.
 DR PDB; 1MET; 15-APR-98.
 DR PDB; 1MEU; 15-APR-98.
 DR PDB; 1OIW; 18-FEB-03.
 DR PDB; 1ODW; 01-APR-97.
 DR PDB; 1OBS; 15-OCT-97.
 DR PDB; 1OBT; 15-OCT-97.
 DR PDB; 1OBU; 15-OCT-97.
 DR PDB; 1RT4; 29-JUL-99.
 DR PDB; 1RT5; 29-JUL-99.
 DR PDB; 1RT7; 29-JUL-99.
 DR PDB; 1RTD; 12-JAN-00.
 DR PDB; 1VRT; 03-APR-96.
 DR PDB; 3PHV; 15-JAN-92.
 DR HIV; K03455; POLSHXB2.
 DR MEROPS; A02.001; -.
 DR InterPro; IPR001969; Asparticase_AS.
 DR InterPro; IPR001037; Integrase_C.
 DR InterPro; IPR003308; Integrase_Zn.
 DR InterPro; IPR009007; Pept_A acid.
 DR InterPro; IPR001995; Peptidase_A2.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR00477; Rvse.
 DR Pfam; PF00552; Integrase_1.
 DR Pfam; PF02022; Integrase_Zn; 1.
 DR Pfam; PF00075; rnaesh; 1.
 DR Pfam; PF00665; rve; 1.
 DR Pfam; PF00077; rvp; 1.
 DR Pfam; PF00078; rvt; 1.
 DR PROSITE; PS00141; ASP_PROTASE; 1.
 DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
 DR AID5; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
 DR Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
 FT CHAIN 57 155
 FT ACT_SITE 81 81
 FT STRAND 58 59
 FT STRAND 66 71
 FT TURN 72 73
 FT STRAND 74 80
 FT TURN 82 83
 FT STRAND 88 89
 FT STRAND 99 104
 FT STRAND 109 122
 FT STRAND 126 133
 FT STRAND 140 141
 FT STRAND 143 146
 FT HELIX 147 150
 FT TURN 152 154
 FT STRAND 154 154
 FT STRAND 167 167

Query Match 37.5%; Score 54; DB 1; Length 1003;

Best Local Similarity 45.2%; Pred. No. 3.3;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKFIGITERG-----RLKEPVHGV 28
 Db 442 KALTEVIPLEAELELANREILKEPVHGV 472
 RESULT 5
 POL_HV10Y
 ID POL_HV10Y STANDARD; PRT; 1003 AA.
 AC P20892;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE POL polypeptide [contains: Protease (Retropepsin) (EC 3.4.23.16);
 DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN POL.
 OS Human immunodeficiency virus type 1 (OYI isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90148544; PubMed=2559749;
 RA Hue T.; Dazza M.C.; Brun-Vezinet F.; Roelants G.E.; Wain-Hobson S.;
 RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
 RT individual presenting an atypical western blot."
 RL AIDS 3:707-715(1989).
 CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
 CC and P1' variable, but often Pro.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphomonoester.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + [DNA] (N).
 CC -1- P1M: Cleavage sites that yield the mature proteins remain to be
 CC determined.
 CC -1- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A
 CC HEALTHY GABONESE INDIVIDUAL.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 CC KNOWN AS THE RETROPEPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M26727; AAA8392.1; -.
 CC HSSP; P03366; IRVL.
 CC HIV; M26727; POLSOYI.
 DR InterPro; IPR001969; Asparticase_AS.
 DR InterPro; IPR001037; Integrase_C.
 DR InterPro; IPR003308; Integrase_Zn.
 DR InterPro; IPR009007; Pept_A acid.
 DR InterPro; IPR001995; Peptidase_A2.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR00477; Rvse.
 DR Pfam; PF00552; Integrase_1.
 DR Pfam; PF02022; Integrase_Zn; 1.
 DR Pfam; PF00075; rnaesh; 1.
 DR Pfam; PF00665; rve; 1.
 DR Pfam; PF00077; rvp; 1.
 DR Pfam; PF00078; rvt; 1.
 DR PROSITE; PS00141; ASP_PROTASE; 1.
 DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
 DR AID5; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
 DR Nuclease; Transferase; RNA-directed DNA polymerase.
 FT CHAIN 57 155
 FT ACT_SITE 81 81
 FT BY SIMILARITY.

SQ SEQUENCE 1003 AA; 113718 MW; APE997A0DEB88A98 CRC64;
 Query Match 37.5%; Score 54; DB 1; Length 1003;
 Best Local Similarity 45.2%; Pred. No. 3.3;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKFIGTTERG-----RIIKEPVHGV 28
 DB 442 KALTEVITPLEBAEELAEENREILKEPVHGV 472
 RESULT 6
 ID POL_HV1Y2 STANDARD; PRT; 1003 AA.
 AC P35963;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
 Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN POL.
 OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=36377;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93021387; PubMed=1404605;
 RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
 Shaw G.M.;
 RT "Complete nucleotide sequence, genome organization, and biological
 properties of human immunodeficiency virus type 1 in vivo: evidence
 for limited defectiveness and complementation.";
 J. Virol. 66:6587-6600(1992).
 CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
 and P1' variable, but often Pro.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 phosphomonoester
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + {DNA} (N).
 CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
 determined.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 KNOWN AS THE RETROPEPSIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M93258; -; NOT_ANNOTATED_CDS.
 DR PIR; B44001; B44001.
 DR PDB; 1K6C; 06-FEB-02.
 DR PDB; 1K6F; 06-FEB-02.
 DR PDB; 1K6T; 06-FEB-02.
 DR PDB; 1K6V; 06-FEB-02.
 DR PDB; 1WJA; 13-MAY-98.
 DR PDB; 1WJC; 13-MAY-98.
 DR MEROPS; A02.001; -;
 DR InterPro; IPR001969; Asparticase_AS.
 DR InterPro; IPR001037; Integrase_C.
 DR InterPro; IPR003308; Integrase_Zn.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001995; Peptidase_A2.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR001584; Rye.
 DR InterPro; IPR000477; RYase.
 DR Pfam; PF00552; Integrase; 1.
 DR Pfam; PF00202; Integrase_Zn; 1.
 DR Pfam; PF00075; RNaseH; 1.
 DR Pfam; PF00665; rve; 1.

DR Pfam; PF00077; rvp; 1.
 DR Pfam; PF00078; rvc; 1.
 DR PROSITE; PS00141; ASD_PROTEASE; 1.
 DR PROSITE; PS50175; ASD_PROT_RETROV; 1.
 KW Aids; Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease;
 RNase; Transferase; RNA-directed DNA polymerase; 3D-structure.
 FT CHAIN 57 155
 FT ACT SITE 81 81
 FT BY SIMILARITY.
 SQ SEQUENCE 1003 AA; 113794 MW; D2F7187FE4E8F49 CRC64;
 Query Match 37.5%; Score 54; DB 1; Length 1003;
 Best Local Similarity 45.2%; Pred. No. 3.3;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKFIGTTERG-----RIIKEPVHGV 28
 DB 442 KALTEVITPLEBAEELAEENREILKEPVHGV 472
 RESULT 7
 ID POL_HV1MN STANDARD; PRT; 1006 AA.
 AC P05961;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 26-FEB-2003 (Rel. 41, Last annotation update)
 DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
 Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN POL.
 OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11696;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88219542; PubMed=3369091;
 RA Garzo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
 RA Farrell K., Wong-Staal F., Gallo R.C., Reltz M.S., Jr.;
 RT "Envelope sequences of two new United States HIV-1 isolates.";
 J. Virol. 64:531-536(1988).
 CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
 and P1' variable, but often Pro.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 phosphomonoester.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + {DNA} (N).
 CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
 determined.
 CC -1- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
 PATIENT IN 1984.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 KNOWN AS THE RETROPEPSIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M17449; -; NOT_ANNOTATED_CDS.
 DR HSP; P03366; IRVL.
 DR HIV; M17449; POLSMN.
 DR MEROPS; A02.001; -;
 DR InterPro; IPR001969; Asparticase_AS.
 DR InterPro; IPR001037; Integrase_C.
 DR InterPro; IPR003308; Integrase_Zn.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001995; Peptidase_A2.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR001584; Rye.
 DR InterPro; IPR000477; RYase.
 DR Pfam; PF00552; Integrase; 1.

FT TURN 135 136
 FT STRAND 137 146
 FT STRAND 152 153
 FT HELIX 155 161
 FT TURN 162 162
 SQ SEQUENCE 1015 AA, 115031 MW, 164702F074A84394 CRC64;
 Query Match 37.5%; Score 54; DB 1; Length 1015;
 Best Local Similarity 45.2%; Pred. No. 3.4;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKFIGITRG-----RIKKEPVHG 28
 DB 454 KALTEVILPTEAELELAENREILKEPVHG 484
 RESULT 11
 POL_HV1PV STANDARD; PRT; 1015 AA.
 AC P03368;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE POL. polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
 DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)]
 GN POL.
 OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 NC NCBI Taxid=11700;
 RN RP SEQUENCE FROM N.A.
 RX MEDLINE=85111157; PubMed=2982104;
 RA Muesing M.A., Smith D.H., Cabral J.L.A., Benton C.V., Laessle L.A.,
 RT AIDS/Lymphadenopathy retrovirus."
 RL Nature 313:450-458(1985).
 CC [2]
 CC REVISION.
 RA Muesing M.A.;
 RL Submitted (XXX-1987) to the HIV data bank.
 CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
 CC and P1' variable, but often Pro.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphonooxester.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + {DNA} (N).
 CC -1- PFM: Cleavage sites that yield the mature proteins remain to be
 CC determined.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 CC KNOWN AS THE RETROPEPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; K02083; AAB59867.1; -;
 CC EMBL; X01762; -; NOT_ANNOTATED_CDS.
 CC PIR; A03967; GNTWVL.
 CC PDB; 1A9M; 17-JUN-98.
 CC PDB; 1A9V; 20-AUG-97.
 CC PDB; 1A9X; 17-SEP-97.
 CC PDB; 1G35; 06-JUN-01.
 CC PDB; 1GMM; 08-NOV-96.
 CC PDB; 1HTG; 31-UTL-94.
 CC PDB; 1HVI; 30-APR-94.
 CC PDB; 1KJH; 06-MAR-92.
 CC PDB; 1NPV; 04-FEB-03.
 CC PDB; 1NPW; 04-FEB-03.

DR HIV; K02083; POL\$PV22.
 DR MEROPS; A02.001; -;
 DR InterPro; IPR001969; Asparticase_AS.
 DR InterPro; IPR001037; Integrase_C.
 DR InterPro; IPR003308; Integrase_Zn.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001995; Peptidase_A2.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR001584; Rve.
 DR InterPro; IPR000477; RNase.
 DR Pfam; PF00552; Integrase; 1.
 DR Pfam; PF02022; Integrase; 1.
 DR Pfam; PF00075; RNaseH; 1.
 DR Pfam; PF00665; Rve; 1.
 DR Pfam; PF00077; Rve; 1.
 DR Pfam; PF00078; Rve; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
 DR KMW; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
 DR Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
 FT CHAIN 69 167
 FT ACT SITE 93 93 BY SIMILARITY.
 SQ SEQUENCE 1015 AA; 115090 MW; 51529D18AC2AF89 CRC64;
 Query Match 37.5%; Score 54; DB 1; Length 1015;
 Best Local Similarity 45.2%; Pred. No. 3.4;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKFIGITRG-----RIKKEPVHG 28
 DB 454 KALTEVILPTEAELELAENREILKEPVHG 484
 RESULT 12
 POL_HV1ND STANDARD; PRT; 1002 AA.
 AC P18802;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE POL. polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
 DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)]
 GN POL.
 OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 NC NCBI Taxid=11695;
 RN RP SEQUENCE FROM N.A.
 RX MEDLINE=90034200; PubMed=2806917;
 RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,
 RA Hampe A., Chermann J.C.;
 RL "Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the
 RL human immunodeficiency virus."
 RT Gene 81:275-284(1989).
 CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
 CC and P1' variable, but often Pro.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphonooxester.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + {DNA} (N).
 CC -1- PFM: Cleavage sites that yield the mature proteins remain to be
 CC determined.
 CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
 CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 CC KNOWN AS THE RETROPEPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M27323; AAA44869.1; -.
CC PIR; J00067; GNJND.
CC HSP; P03366; IHMV.
CC HIV; M27323; POL$NDK.
CC MEROPS; A02.001; -.
CC InterPro; IPR001969; Asparticase_AS.
CC InterPro; IPR001037; Integrase_C.
CC InterPro; IPR00308; Integrase_Zn.
CC InterPro; IPR009007; Pept_A_acid.
CC InterPro; IPR001995; Peptidase_A2.
CC InterPro; IPR002156; RNaseH.
CC InterPro; IPR001584; Rve.
CC InterPro; IPR000477; RVse.
CC Pfam; PF00552; Integrase_1.
CC Pfam; PF02022; Integrase_Zn; 1.
CC Pfam; PF00075; rnaseH; 1.
CC Pfam; PF00665; rve; 1.
CC Pfam; PF00077; rvp; 1.
CC Pfam; PF00078; rvt; 1.
CC PROSITE; PS00141; ASP_PROTEASE; 1.
CC PROSITE; PS50175; ASP_PROT_RETROV; 1.
CC AIDS; Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease;
CC Nuclease; Transferase; RNA-directed DNA polymerase.
CC CHAIN 56 154
CC ACT SITE 80 80 BY SIMILARITY.
CC SEQUENCE 1002 AA; 113621 MW; 58D868C9896CC1 CRC64;
SQ
Query Match 36.8%; Score 53; DB 1; Length 1002;
Best Local Similarity 41.9%; Pred. No. 4.7;
Matches 13; Conservative 4; Mismatches 6; Indels 8; Gaps 1;
Qy 6 KANSKFIGITERG-----RLKEPVHGV 28
Db 441 KALTEVPLTEAELELAENREILKEPVHGV 471
RESULT 13
POL_HY1N5 STANDARD; PRT; 1003 AA.
ID POL_HY1N5
AC P12497;
DC 01-OCT-1989 (Rel. 12; Created)
DT 01-OCT-1989 (Rel. 12; Last sequence update)
DT 10-OCT-2003 (Rel. 42; Last annotation update)
DE POL polypeptide (Contains: Protease (Retropesin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)]].
GN POL.
OS Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11698;
RN [1]
RP SEQUENCE FROM N.A. (CLONE PML4-3).
RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;
RL Submitted (JUN-1988) to the EMBL/Genbank/DBD databases.
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 57-155.
RX MEDLINE=90354401; PubMed=2201682;
RA Fitzgerald P.M.D., McKeever B.M., van Middlesworth J.F.,
RA Springer J.P., Heimbach J.C., Liu C.-T., Herber W.K., Dixon R.A.F.,
RA Darke P.L.;
RT "Crystallographic analysis of a complex between human
RT immunodeficiency virus type 1 protease and acetyl-peptatin at 2.0-A
RT resolution.";
RU J. Biol. Chem. 265:14209-14219(1990).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
CC and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester..
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
```

```
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M1921; AAA44968.1; -.
CC PDB; 5HP; 15-OCT-91.
CC PDB; 4PHV; 31-OCT-93.
CC PDB; 1B9F; 19-JUL-99.
CC PDB; 1BHL; 28-OCT-98.
CC PDB; 1B14; 18-NOV-98.
CC PDB; 1B1S; 16-SEP-98.
CC PDB; 1B1U; 16-SEP-98.
CC PDB; 1WJB; 13-MAY-98.
CC PDB; 1WJD; 13-MAY-98.
CC PDB; 21NG; 12-MAR-97.
CC PDB; 9HP; 15-JUL-92.
CC HIV; M1921; POL$ML43.
CC MEROPS; A02.001; -.
CC InterPro; IPR001969; Asparticase_AS.
CC InterPro; IPR001037; Integrase_C.
CC InterPro; IPR00308; Integrase_Zn.
CC InterPro; IPR009007; Pept_A_acid.
CC InterPro; IPR001995; Peptidase_A2.
CC InterPro; IPR002156; RNaseH.
CC InterPro; IPR001584; Rve.
CC InterPro; IPR000477; RVse.
CC Pfam; PF00552; Integrase_1.
CC Pfam; PF02022; Integrase_Zn; 1.
CC Pfam; PF00075; rnaseH; 1.
CC Pfam; PF00665; rve; 1.
CC Pfam; PF00077; rvp; 1.
CC Pfam; PF00078; rvt; 1.
CC PROSITE; PS00141; ASP_PROTEASE; 1.
CC PROSITE; PS50175; ASP_PROT_RETROV; 1.
CC AIDS; Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease;
CC Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
CC CHAIN 57 155
CC ACT SITE 81 81
CC SEQUENCE 1003 AA; 113535 MW; EB9CDBA30E114A55 CRC64;
SQ
Query Match 36.8%; Score 53; DB 1; Length 1003;
Best Local Similarity 41.9%; Pred. No. 4.8;
Matches 13; Conservative 4; Mismatches 6; Indels 8; Gaps 1;
Qy 6 KANSKFIGITERG-----RLKEPVHGV 28
Db 442 KALTEVPLTEAELELAENREILKEPVHGV 472
RESULT 14
```

```
POL_HV1MA
ID POL_HV1MA STANDARD; PRT; 1002 AA.
AC P04588; Q79582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1988 (Rel. 09, Last annotation update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (MNL isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11697;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients.";
RL Cell 46:63-74(1986).
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ (DNA) (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; X0415; CA28012.1; -.
CC EMBL; A07116; CA00619.1; -.
CC PDB; 1HHJ; 31-OCT-93.
CC HIV; K03456; POLSML.
CC MEROPS; A02.001; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROT_PROTASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
KW AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 56 154
FT ACT_SITE 80 80 BY SIMILARITY.
SQ SEQUENCE 1002 AA; 113537 MW; 569A29D69AC6AC5 CRC64;
Query Match 36.1%; Score 52; DB 1; Length 1002;
Best Local Similarity 66.7%; Pred. No. 6.8;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
RESULT 15
ID POL_HV1RH STANDARD; PRT; 1002 AA.
AC P05959;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (RF/HAT isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11701;
RN [1]
RP SEQUENCE FROM N.A.
RA Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S.,
RA Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C.,
RA Wong-Staal F.;
RT Submitted (XXX-1987) to the HIV data bank.
RL
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ (DNA) (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; M17451; AAA5053.1; -.
CC HSSP; P04585; 1RTH.
CC HIV; M17451; POLSRF.
CC MEROPS; A02.001; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROT_PROTASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
KW AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
Nuclease; Transferase; RNA-directed DNA polymerase.
FT CHAIN 56 154
FT ACT_SITE 80 80 BY SIMILARITY.
SQ SEQUENCE 1002 AA; 113755 MW; 864341718E5C48C2 CRC64;
Query Match 36.1%; Score 52; DB 1; Length 1002;
Best Local Similarity 66.7%; Pred. No. 6.8;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

Tue Jul 20 07:06:41 2004

us-09-673-166b-274.rsp

Page 13

Search completed: July 20, 2004, 06:31:01
Job time : 6.5082 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:43 ; Search time 28.2295 Seconds

(without alignments)
312.953 Million cell updates/sec

Title: US-09-673-166B-274

Perfect score: 144

Sequence: 1 GRQYKANSKFIQITERGRILKEPVHGV 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacterioplasmid:*
17: sp_archaeplastid:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	48.6	1310	2	Q93N27
2	60	41.7	1427	15	Q9WF98
3	57	39.6	1003	15	Q79792
4	57	39.6	1003	15	P89972
5	56	38.9	322	15	Q7SKY9
6	56	38.9	323	15	Q7SKY2
7	56	38.9	420	15	Q8Q736
8	56	38.9	420	15	Q8Q761
9	56	38.9	423	15	Q8AF11
10	56	38.9	524	15	Q9IDB5
11	56	38.9	560	15	Q71151
12	56	38.9	1433	15	Q71151
13	55	38.2	364	15	Q8ADE1
14	55	38.2	420	15	P90232
15	55	38.2	423	15	Q8Q784
16	55	38.2	423	15	Q8JDC4
					Q8AHG3

17	55	38.2	423	15	Q8AH44	Q8AH44 human immun
18	55	38.2	423	15	Q8AGZ4	Q8AGZ4 human immun
19	55	38.2	434	15	Q7ZPA1	Q7ZPA1 human immun
20	55	38.2	434	15	Q7ZP95	Q7ZP95 human immun
21	55	38.2	555	15	Q9J076	Q9J076 human immun
22	55	38.2	564	15	Q9J4V3	Q9J4V3 human immun
23	55	38.2	564	15	Q9J4V2	Q9J4V2 human immun
24	55	38.2	564	15	Q9J4V1	Q9J4V1 human immun
25	55	38.2	593	15	Q8Q876	Q8Q876 human immun
26	55	38.2	1000	15	Q8AUC5	Q8AUC5 human immun
27	55	38.2	1005	15	Q8AFB1	Q8AFB1 human immun
28	55	38.2	1429	15	Q9WF84	Q9WF84 human immun
29	54	37.5	237	15	Q9WGV3	Q9WGV3 human immun
30	54	37.5	237	15	Q9WGV7	Q9WGV7 human immun
31	54	37.5	237	15	Q9WGV5	Q9WGV5 human immun
32	54	37.5	237	15	Q9WGV4	Q9WGV4 human immun
33	54	37.5	237	15	Q9WGV6	Q9WGV6 human immun
34	54	37.5	237	15	Q9WGV3	Q9WGV3 human immun
35	54	37.5	237	15	Q9WGV4	Q9WGV4 human immun
36	54	37.5	237	15	Q9WGV2	Q9WGV2 human immun
37	54	37.5	237	15	Q9WGV5	Q9WGV5 human immun
38	54	37.5	237	15	Q9WGV6	Q9WGV6 human immun
39	54	37.5	237	15	Q9WGV3	Q9WGV3 human immun
40	54	37.5	237	15	Q9WGV4	Q9WGV4 human immun
41	54	37.5	237	15	Q9WGV6	Q9WGV6 human immun
42	54	37.5	237	15	Q9WGV3	Q9WGV3 human immun
43	54	37.5	257	15	Q9WGV4	Q9WGV4 human immun
44	54	37.5	319	15	Q7SKW0	Q7SKW0 human immun
45	54	37.5	321	15	Q9JMW5	Q9JMW5 human immun

ALIGNMENTS

RESULT 1

ID Q93N27 PRELIMINARY; PRT; 1310 AA.
AC Q93N27;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Tetanus toxin (Fragment).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxId=1513;
RN [1]
RP SEQUENCE FROM N.A.
RA Shumin Z., Dianliang L.,
RT "Cloning and sequence analysis of tetanus toxin gene.",
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF389424; AAK72964.2; -
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0009405; P:proteolysis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008985; ConA_like_lect_g1.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR001064; Kunitz_legume.
DR InterPro; IPR000395; Peptidase_M27.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BOWTOXITXIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON_TER 1
FT NON_TER 1310
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418B450 CRC64;

Query Match

48.6%; Score 70; DB 2; Length 1310;

Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 OYKANSKFIGITE 16
831 OYKANSKFIGITE 844
Db

RESULT 2
Q9WF98 PRELIMINARY; PRT; 1427 AA.
ID Q9WF98
AC Q9WF98
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Gag-pol polyprotein (Gag polyprotein).
GN GAG.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_Taxid=11676;
OX NCBI_Taxid=11676;
RN SEQUENCE FROM N.A.
RC STRAIN=C-96BMD14;
RX MEDLINE=99214383; PubMed=10196340;
RA Novitsky V.A., Montano M.A., McLane M.F., Renjifo B., Vanberg F.,
RA Foley B.T., Ndung'u T.P., Rahman M., Makheima M.O., Marlink R.,
RA Essex M.;
RT "Molecular cloning and phylogenetic analysis of human immunodeficiency
RT virus type 1 subtype C: a set of 23 full-length clones from
RT Botswana.";
RT J. Virol. 73:4427-4432(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C-96BMD14;
RA Novitsky V.A., Montano M.A., McLane M.F., Renjifo B., Vanberg F.,
RA Foley B.T., Ndung'u T.P., Marlink R., Essex M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS
CC (BY SIMILARITY).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC EMBL: AF110977; AAD17153.1; -;
CC PIR: S49086; S49086.
DR HSP: P03366; IHNV.
DR GO: GO:0019012; C:virion; IEA.
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0004519; F:endonuclease activity; IEA.
DR GO: GO:0008907; F:integrase activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0004523; F:ribonuclease H activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR GO: GO:0015074; P:DNA integration; IEA.
DR GO: GO:0006310; P:DNA recombination; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR GO: GO:0016032; P:viral life cycle; IEA.
DR InterPro: IPR001969; Asprotease_AS.
DR InterPro: IPR000721; Gag_P24.
DR InterPro: IPR001037; Integrase_C.
DR InterPro: IPR003308; Integrase_Zn.
DR InterPro: IPR001995; Peptidase_A2.
DR InterPro: IPR000907; Pept_Acid.
DR InterPro: IPR000071; Retrov_P17.
DR InterPro: IPR008916; Retrov_Capsid_C.
DR InterPro: IPR008919; Retrov_Capsid_N.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR001584; Rve.
DR InterPro: IPR000477; RTase.

DR InterPro: IPR001878; Znf_CCHC.
DR Pfam: PF00540; Gag_P17; 1.
DR Pfam: PF00607; Gag_P24; 1.
DR Pfam: PF00552; Integrase; 1.
DR Pfam: PF02022; Integrase_Zn; 1.
DR Pfam: PF00075; rnaseh; 1.
DR Pfam: PF00655; rve; 1.
DR Pfam: PF00077; rvp; 1.
DR Pfam: PF00078; rvc; 1.
DR Pfam: PF00939; zF_CCHC; 2.
DR PRINTS: PR00939; C2HCNFRINGER.
DR PRINTS: PR00234; HIVMATRIX.
DR SMART: SM00343; ZNF_C2HC; 2.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
DR PROSITE: PS50158; ZF_CCHC; 2.
DR AIDS; Aspartyl protease; Core protein; Endonuclease; Hydrolyase;
KW Nucleotidyltransferase; Polyprotein; Protease;
KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1427 AA; 160409 MW; E38F85FC6A0B369 CRC64;
Query Match 41.7%; Score 60; DB 15; Length 1427;
Best Local Similarity 73.3%; Pred. No. 5.5;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 14 ITERGRIILKEPVHGV 28
882 LAENGRIILKEPVHGV 896
Db

RESULT 3
Q79792 PRELIMINARY; PRT; 1003 AA.
ID Q79792
AC Q79792
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE pol protein (Fragment).
GN GAG-POL OR POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_Taxid=11676;
OX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NY5.
RX MEDLINE=86259728; PubMed=3014529;
RA Willey R.L., Rutledge R.A., Dias S., Folks T., Theodore T.,
RA Buckler C.E., Martin M.A.;
RT Identification of acquired immunodeficiency syndrome retrovirus.";
RT gene of the acquired immunodeficiency syndrome retrovirus.";
CC Proc. Natl. Acad. Sci. U.S.A. 83:5038-5042(1986).
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS
CC (BY SIMILARITY).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC EMBL: M38431; AAB04037.1; -;
CC PIR: A47330; A47330.
CC PIR: B47330; B47330.
CC PIR: C47330; C47330.
CC PIR: D47330; D47330.
CC PIR: E47330; E47330.
CC PIR: F47330; F47330.
CC PIR: S32058; S32058.
CC PIR: S32132; S32132.
CC PIR: S32140; S32140.
DR HSP: P12497; SHVP.
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0004519; F:endonuclease activity; IEA.
DR GO: GO:0008907; F:integrase activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0004523; F:ribonuclease H activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.

```
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Asparticase_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVise.
DR Pfam; PF02052; Integrase_1.
DR Pfam; PF02022; Integrase_Zn_1.
DR Pfam; PF00075; RNaseH_1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Endonuclease; Hydrolyase; Nucleotidyltransferase;
KW Polyprotein; Protease; RNA-directed DNA polymerase; Transferase.
FT NON_TER
SQ SEQUENCE 1003 AA; 113729 MW; E24A73675C1651A1 CRC64;

Query Match          39.6%; Score 57; DB 15; Length 1003;
Best Local Similarity 43.8%; Pred. No. 11;
Matches 14; Conservative 4; Mismatches 6; Indels 8; Gaps 1;

QY 5 KANSKFIGITRG-----RIIKEPVHG 28
DB 441 KALTEVPLTEAELELAENREILKEPVHG 472
|||:::|||||
P89972 PRELIMINARY; PRT; 1003 AA.
ID P89972;
AC P89972;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pol polyprotein (Fragment).
GN GAG-POL OR POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Ghosh S., Ghosh E., Shaw G.M., Hahn B.;
RT "Human Immunodeficiency Virus type 1 complete genomic sequence.";
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA McCluskey K.A.;
RT Submitted (FEB-1995) to the EMBL/Genbank/DBJ databases.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS
CC (BY SIMILARITY).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL; U21135; AAC32294.1; -.
DR PIR; A47330; A47330.
DR PIR; B47330; B47330.
DR PIR; C47330; C47330.
DR PIR; D47330; D47330.
DR PIR; E47330; E47330.
DR PIR; F47330; F47330.
DR PIR; S32132; S32132.
DR HSSP; P04585; 1RTH.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
```

```
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0008907; F:integrase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA recombination; IEA.
DR GO; GO:0006310; P:DNA integration; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Asparticase_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVise.
DR Pfam; PF02052; Integrase_1.
DR Pfam; PF02022; Integrase_Zn_1.
DR Pfam; PF00075; RNaseH_1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Endonuclease; Hydrolyase; Nucleotidyltransferase;
KW Polyprotein; Protease; RNA-directed DNA polymerase; Transferase.
FT NON_TER
SQ SEQUENCE 1003 AA; 113833 MW; 29A57DAF295477A9 CRC64;

Query Match          39.6%; Score 57; DB 15; Length 1003;
Best Local Similarity 48.4%; Pred. No. 11;
Matches 15; Conservative 2; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITRG-----RIIKEPVHG 28
DB 442 KALTEIIPITBEAELELAENREILKEPVHG 472
|||:::|||||
Q7SKY9 PRELIMINARY; PRT; 322 AA.
ID Q7SKY9;
AC Q7SKY9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Reverse transcriptase (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3999-7303.4;
RA Gonzales M.J., Johnston E., Dupnik K.M., Imamichi T., Shafer R.W.;
RT "Coincidence of Reverse Transcriptase Inhibitor Resistance Mutations
RT Detected by Population-based Sequencing.";
RL Submitted (JUL-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY351751; AAQ18859.1; -.
DR RNA-directed DNA polymerase.
FT NON_TER
FT NON_TER
SQ SEQUENCE 322 AA; 37354 MW; 095623DBAC3D0D5 CRC64;

Query Match          38.9%; Score 56; DB 15; Length 322;
Best Local Similarity 45.2%; Pred. No. 4.4;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITRG-----RIIKEPVHG 28
|||:::|||||
```

Db 287 KALTEVILSTEAELELAENREILKEPVHGV 317

RESULT 6

Q7SKZ2 PRELIMINARY; PRT; 323 AA.

AC Q7SKZ2; 01-OCT-2003 (TReMBLrel. 25, Created)

DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE Reverse transcriptase (Fragment).

GN POL.

OS Human immunodeficiency virus 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OC NCBI_TaxID=11676;

RN [1] TaxID=11676;

RP SEQUENCE FROM N.A.

RC STRAIN=399-7303.1;

RA Gonzales M.J., Johnston E., Dupnik K.M., Imachi T., Shafer R.W.;

RT "Colinearity of Reverse Transcriptase Inhibitor Resistance Mutations Detected by Population-Based Sequencing."

RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY51748; AA018856.1; "

DT EMBL-directed DNA polymerase.

FT NON_TER 1 1

FT NON_TER 323 323

SEQ SEQUENCE 323 AA; 37404 MW; 59BF3FE314FB3B93 CRC64;

Query Match 38.9%; Score 56; DB 15; Length 323;

Best Local Similarity 45.2%; Pred. No. 4.4;

Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITERG-----RIKKEPVHGV 28

Db 287 KALTEVILSTEAELELAENREILKEPVHGV 317

RESULT 7

Q8Q736 PRELIMINARY; PRT; 420 AA.

AC Q8Q736; 01-JUN-2002 (TReMBLrel. 21, Created)

DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE Pol protein (Fragment).

GN POL.

OS Human immunodeficiency virus 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OC NCBI_TaxID=11676;

RN [1] TaxID=11676;

RP SEQUENCE FROM N.A.

RC STRAIN=537;

RA Bahleman S.H., Becker-Pergola G., Deseyve M., Macra M., Guay L.A.,

RT "Impact of HIV-1 subtype on women receiving single dose NVP prophylaxis to prevent HIV-1 vertical transmission (HIVNET012).";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO KNOWN AS THE RETROPEPSIN FAMILY.

DR EMBL; AF388147; AAL84089.1; "

DR HSSP; P03366; 3HVT.

FT NON_TER 420 420

SEQ SEQUENCE 420 AA; 47981 MW; E23956F9A87B57C6 CRC64;

Query Match 38.9%; Score 56; DB 15; Length 420;

Best Local Similarity 40.6%; Pred. No. 5.9;

Matches 13; Conservative 5; Mismatches 6; Indels 8; Gaps 1;

QY 5 IKANSKFIGITERG-----RIKKEPVHGV 28

Db 385 VKALTEVILSTEAELELAENREILKEPVHGV 416

RESULT 9

Dr Pfam; PF00078; rvt; 1.

DR PROSITE; PS00141; ASP_PROTEASE; 1.

DR PROSITE; PS01175; ASP_PROT_RETROV; 1.

KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;

KW Transferase.

FT NON_TER 1 1

FT NON_TER 420 420

SEQ SEQUENCE 420 AA; 47880 MW; D65393BC59P921B0 CRC64;

Query Match 38.9%; Score 56; DB 15; Length 420;

Best Local Similarity 40.6%; Pred. No. 5.9;

Matches 13; Conservative 5; Mismatches 6; Indels 8; Gaps 1;

QY 5 IKANSKFIGITERG-----RIKKEPVHGV 28

Db 385 VKALTEVILSTEAELELAENREILKEPVHGV 416

RESULT 8

Q8Q761 PRELIMINARY; PRT; 420 AA.

AC Q8Q761; 01-JUN-2002 (TReMBLrel. 21, Created)

DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE Pol protein (Fragment).

GN POL.

OS Human immunodeficiency virus 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OC NCBI_TaxID=11676;

RN [1] TaxID=11676;

RP SEQUENCE FROM N.A.

RC STRAIN=751;

RA Bahleman S.H., Becker-Pergola G., Deseyve M., Macra M., Guay L.A.,

RT "Impact of HIV-1 subtype on women receiving single dose NVP prophylaxis to prevent HIV-1 vertical transmission (HIVNET012).";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO KNOWN AS THE RETROPEPSIN FAMILY.

DR EMBL; AF388122; AAL84064.1; "

DR HSSP; P03366; 2HMI.

FT NON_TER 420 420

SEQ SEQUENCE 420 AA; 47981 MW; E23956F9A87B57C6 CRC64;

Query Match 38.9%; Score 56; DB 15; Length 420;

Best Local Similarity 40.6%; Pred. No. 5.9;

Matches 13; Conservative 5; Mismatches 6; Indels 8; Gaps 1;

QY 5 IKANSKFIGITERG-----RIKKEPVHGV 28

Db 385 VKALTEVILSTEAELELAENREILKEPVHGV 416

RESULT 9

```

Q8AFI.1
ID 08AFI.1 PRELIMINARY; PRT; 423 AA.
AC 08AFI.1;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=lv904;
RA Brindeiro P.A., Brindeiro R.M., Mortensen C., Hertogs K., de Vroey V.,
RA Tanuti N.P.M., Sion F.S., de Sa C.A.M., Machado D.M., Succi R.C.M.,
RA Tanuti A.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY145784; AA012458.1; -;
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Asprotease AS.
DR InterPro; IPR009007; Peptidase A2.
DR InterPro; IPR000477; RVse.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 423 AA; 48130 MW; 32F57BF19B068F1 CRC64;

Query Match
Best Local Similarity 45.2%; Score 56; DB 15; Length 423;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITRG-----RIKPEVHG 28
Db 386 KALTEVISLTERAELELAENREILKEPVHG 416

RESULT 10
Q9IDB5 PRELIMINARY; PRT; 524 AA.
AC Q9IDB5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE POL (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=FROM N.A.
RC STRAIN=99ER-WP1321;
RX MEDLINE=20514217; PubMed=11060045;
RX Verge L., Peters M., Mpoudi-Ekoule E., Bourgeois A., Liegeois F.,
RX Toure-Kane C., Moup S., Mulanga-Kabeya C., Saman E., Jourdan J.,
RX Reynes J., Delaporte E.;
RT "Genetic Diversity of Protease and Reverse Transcriptase Sequences in
RT non-subtype-b human immunodeficiency virus type 1 strains: evidence of
RT many minor drug resistance mutations in treatment-naïve patients.";
RL J. Clin. Microbiol. 38:3919-3925(2000).
CC -1- STIMULATORY. THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPESTIN FAMILY.
DR EMBL; AJ287045; CAC03087.1; -;
DR PIR; A47330; A47330.
DR PIR; C47330; C47330.

```

```

DR PIR; D47330; D47330.
DR PIR; F47330; F47330.
DR HSSP; P04585; 1RTH.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Asprotease AS.
DR InterPro; IPR001995; Peptidase A2.
DR InterPro; IPR009007; Peptidase A2.
DR InterPro; IPR000477; RVse.
DR Pfam; PF00078; rvt; 1.
DR Pfam; PF00077; rvp; 1.
DR PROSITE; PS00141; ASP_PROTASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW Transferase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 524 AA; 60460 MW; 8780BEA11696EB7C CRC64;

Query Match
Best Local Similarity 45.2%; Score 56; DB 15; Length 524;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITRG-----RIKPEVHG 28
Db 389 KALTEVISLTERAELELAENREILKEPVHG 419

RESULT 11
Q71151 PRELIMINARY; PRT; 560 AA.
AC Q71151;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Reverse transcriptase (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=VE4;
RC STRAIN=VE4;
RX MEDLINE=96093896; PubMed=7576917;
RX Quinones-Mateu M.E., Holguin A., Dopazo J., Rota T.R., Domingo E.;
RT "Molecular characterization of human immunodeficiency virus type 1
RT isolates from Venezuela.";
RL AIDS Res. Hum. Retroviruses 11:605-616(1995).
RN [2]
RN [2]
RC STRAIN=FROM N.A.
RC STRAIN=VE4;
RX MEDLINE=97000986; PubMed=8844016;
RX Quinones-Mateu M.E., Holguin A., Dopazo J., Najera I., Domingo E.;
RT "Point mutant frequencies in the pol gene of human immunodeficiency
RT virus type 1 are two- to threefold lower than those of env.";
RL AIDS Res. Hum. Retroviruses 12:1117-1128(1996).
DR EMBL; U16771; AAC55688.1; -;
DR PIR; A47330; A47330.
DR PIR; B47330; B47330.
DR PIR; C47330; C47330.
DR PIR; D47330; D47330.
DR PIR; F47330; F47330.
DR PIR; S32132; S32132.
DR PIR; S32140; S32140.
DR HSSP; P03366; 1HRH.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.

```

DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006278; F:RNA dependent DNA replication; IEA.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF00075; rnaseh; 1.
 DR Pfam; PF00078; rvt; 1.
 DR RNA-directed DNA polymerase; Transferase.
 FT NON_TER 1
 FT NON_TER 560
 SQ SEQUENCE 560 AA; 64354 MW; 8C29E9B280090C8F CRC64;
 Query Match 38.9%; Score 56; DB 15; Length 560;
 Best Local Similarity 45.2%; Pred. No. 8.1;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKPIGTERG-----RIKKEPVHGV 28
 Db 287 KALTEVISTEEAELELANRELIKEPVHGV 317
 RESULT 12
 Q8ADE1 PRELIMINARY; PRT; 1433 AA.
 AC Q8ADE1;
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 25, Last annotation update)
 DE Gag-pol fusion polypeptide (Fragment).
 GN GAG-POL.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=98UG57143;
 RX MEDLINE=20564795; PubMed=11112486;
 RA Salminen M.O., Ehrenberg P.K., Mascola J.R., Dayhoff D.E., Merling R.,
 Blake B., Lander M., Hegerich S., Polonis V.R., Birx D.L., Robb M.L.,
 McCutchan F.E., Michael N.L.;
 RT "Construction and biological characterization of infectious molecular
 clones of HIV-1 subtypes B and E (CRF01_AB) generated by the
 polymerase chain reaction.";
 RT Virology 278:103-110(2000).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=98UG57143;
 RX MEDLINE=22375625; PubMed=12487816;
 RA Harris M.E., Serwadda D., Sewankambo N., Mabweire F., Kim B.,
 Kigozi G., Kiwanuka N., Phillips J.B., Meehan M., Lutalo T.,
 Lane J.R., Merling R., Gray R., Mawer M., Birx D.L., Robb M.L.,
 McCutchan F.E.;
 RT "Among 46 Near Full Length HIV Type 1 Genome Sequences from Rakai
 District, Uganda, Subtype D and AD Recombinants Predominate.";
 RT AIDS Res. Hum. Retroviruses 18:1281-1290(2002).
 RL EMBL: AF484514; AAN73763.1; -
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008907; F:integrase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0015074; P:DNA integration; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
 DR GO; GO:0016032; P:viral life cycle; IEA.
 DR InterPro; IPR001969; Aspartase_AS.
 DR InterPro; IPR000721; Gag_P24.
 DR InterPro; IPR001037; Integrase_C.
 DR InterPro; IPR003308; Integrase_Zn.
 DR InterPro; IPR001995; Peptidase_A2.

DR InterPro; IPR009007; Pept_A acid.
 DR InterPro; IPR000071; Retrovir_p17.
 DR InterPro; IPR008816; Retrov_capsid_C.
 DR InterPro; IPR008919; Retrov_capsid_N.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR001584; Rve.
 DR InterPro; IPR000477; RVTse.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00540; Gag_p17; 1.
 DR Pfam; PF00607; Gag_P24; 1.
 DR Pfam; PF00552; Integrase; 1.
 DR Pfam; PF02022; Integrase_Zn; 1.
 DR Pfam; PF00075; rnaseh; 1.
 DR Pfam; PF00655; rve; 1.
 DR Pfam; PF00077; rvt; 1.
 DR Pfam; PF00078; rvt; 1.
 DR Pfam; PF00098; zf_CCHC; 2.
 DR PRINTS; PR00939; C2HCZNFINGER.
 DR PRINTS; PR00234; HIVMATRIX.
 DR SMART; SM00343; ZNF_C2HC; 2.
 DR PROSITE; PS00141; ASP_PROTASE; 1.
 DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
 DR PROSITE; PS50158; ZF_CCHC; 2.
 KW Polypeptide.
 FT NON_TER 1
 FT NON_TER 161645
 SQ SEQUENCE 1433 AA; 161645 MW; EB213A9BCB7D583 CRC64;
 Query Match 38.9%; Score 56; DB 15; Length 1433;
 Best Local Similarity 45.2%; Pred. No. 22;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKPIGTERG-----RIKKEPVHGV 28
 Db 872 KALTEVISTEEAELELANRELIKEPVHGV 902
 RESULT 13
 P90232 PRELIMINARY; PRT; 364 AA.
 ID P90232;
 AC P90232;
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE POL polypeptide (Fragment).
 GN POL.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IND 4;
 RX MEDLINE=97104208; PubMed=8948366;
 RA Soto-Ramirez L.B., Tripathy S., Remjifo B., Essex M.;
 RT "HIV-1 pol sequences from India fit distinct subtype pattern.";
 RT J. Acquir. Immune Defic. Syndr. Hum. Retrovir. 13:299-307(1996).
 RL EMBL: U31242; AAB40987.1; -
 DR HSSP; P03366; IDLO.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF00078; rvt; 1.
 DR Polypeptide; RNA-directed DNA polymerase; Transferase.
 FT NON_TER 1
 FT NON_TER 1
 FT CHAIN 8
 FT CHAIN 364
 FT NON_TER 364
 FT NON_TER 42031
 SQ SEQUENCE 364 AA; 42031 MW; F4F0479B008590AF CRC64;
 Query Match 38.2%; Score 55; DB 15; Length 364;
 Best Local Similarity 40.6%; Pred. No. 7.2;
 Matches 13; Conservative 4; Mismatches 7; Indels 8; Gaps 1;

```

QY      5 IKANSKFIGITRG-----RIIKEPVHGV 28
      || : : ||| ||| |||
DB      293 KALTEVSVLTERAEALAEENRIIKEPVHGV 324

RESULT 14
Q80784      PRELIMINARY;      PRT;      420 AA.
ID      O80784;
AC      O80784;
DT      01-OCT-2002 (TRENBLrel. 21, Created)
DT      01-OCT-2002 (TRENBLrel. 21, Last sequence update)
DE      01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE      Pol protein (Fragment).
GN      POL.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX      NCBI_Taxid=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=607;
RA      Esheleman S.H., Becker-Pergola G., Deseve M., Macna M., Guay L.A.,
RT      Cunningham S., Musoke P., Mmlro F., Jackson J.B.;
RT      "Impact of HIV-1 subtype on women receiving single dose NVP
RT      prophylaxis to prevent HIV-1 vertical transmission (HIVNET012).";
RL      Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC      -! SIMILARITY: THE PROTEASE BELONGS TO FBPTIDASE FAMILY A2; ALSO
      KNOWN AS THE RETROPEPSIN FAMILY.
DR      EMBL; AF388099; AAL84041.1; -.
DR      HSSP; P03366; 2HM1.
DR      GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR      GO; GO:0008233; F:peptidase activity; IEA.
DR      GO; GO:0003723; F:RNA binding; IEA.
DR      GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR      GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR      InterPro: IPR001969; Aspprotease AS.
DR      InterPro: IPR009007; Peptidase A2.
DR      InterPro: IPR000477; RVTse.
DR      Pfam; PF00077; rvp; 1.
DR      Pfam; PF00078; rvt; 1.
DR      PROSITE; PS00141; ASP_PROTEASE; 1.
DR      PROSITE; PS50175; ASP_PROT_RETROV; 1.
KW      Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
      Transierase.
FT      NON_TER      1
FT      NON_TER      420
SQ      SEQUENCE      420 AA; 48035 MW; E871192B0C277E80 CRC64;

Query Match      38.2%; Score 55; DB 15; Length 420;
Best Local Similarity 41.9%; Pred. No. 8.4;
Matches 13; Conservative 4; Mismatches 6; Indels 8; Gaps 1;

QY      6 KANSKFIGITRG-----RIIKEPVHGV 28
      || : : ||| ||| |||
DB      386 KALTEVSVLTERAEALAEENRIIKEPVHGV 416

RESULT 15
Q8JDC4      PRELIMINARY;      PRT;      423 AA.
ID      O8JDC4;
AC      O8JDC4;
DT      01-OCT-2002 (TRENBLrel. 22, Created)
DT      01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DE      01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE      Pol protein (Fragment).
GN      POL.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX      NCBI_Taxid=11676;
RN      [1]
RP      SEQUENCE FROM N.A.

```

```

RC      STRAIN=148154F;
RA      Macna M.R., Esheleman S.H.;
RT      "Impact of maternal HIV-1 subtype on Ugandan infants (HIVNET 012).";
RL      Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF410250; AMN02560.1; -.
DR      GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR      GO; GO:0003723; F:RNA binding; IEA.
DR      GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR      GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR      InterPro: IPR001969; Aspprotease AS.
DR      InterPro: IPR001969; Aspprotease AS.
DR      InterPro: IPR009007; Peptidase A2.
DR      InterPro: IPR000477; RVTse.
DR      Pfam; PF00077; rvp; 1.
DR      Pfam; PF00078; rvt; 1.
DR      PROSITE; PS00141; ASP_PROTEASE; 1.
DR      PROSITE; PS50175; ASP_PROT_RETROV; 1.
KW      RNA-directed DNA polymerase; Transferase.
FT      NON_TER      1
FT      NON_TER      423
SQ      SEQUENCE      423 AA; 48265 MW; FEF645A37D43C405 CRC64;

Query Match      38.2%; Score 55; DB 15; Length 423;
Best Local Similarity 41.9%; Pred. No. 8.4;
Matches 13; Conservative 4; Mismatches 6; Indels 8; Gaps 1;

QY      6 KANSKFIGITRG-----RIIKEPVHGV 28
      || : : ||| ||| |||
DB      386 KALTEVSVLTERAEALAEENRIIKEPVHGV 416

```

Search completed: July 20, 2004, 06:33:16
 Job time : 29.2295 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:07:17 ; Search time 41.7705 Seconds
(without alignments)
189.400 Million cell updates/sec

Title: US-09-673-166B-274

Perfect score: 144

Sequence: 1 GRQYTKANSKFTGTERGRILKEPVHGV 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	111.5	77.4	25	4	AA884738
2	97	67.4	28	2	AA53552
3	86	59.7	29	2	AA53551
4	81	56.2	29	5	AA879182
5	79	54.9	216	3	AA92665
6	78.5	54.5	31	3	AA82632
7	78	54.2	750	3	AA92639
8	77	53.5	37	2	AA65389
9	76	52.8	29	2	AA83561
10	76	52.8	37	2	AA65383
11	76	52.8	47	2	AA62723
12	76	52.8	750	3	AA92638
13	75	52.1	25	3	AA92652
14	75	52.1	25	3	AA92651
15	75	52.1	31	5	AA879179
16	75	52.1	50	2	AAW06132
17	75	52.1	50	2	AAW46447
18	75	52.1	109	4	AA820145
19	75	52.1	158	2	AA881329
20	75	52.1	158	2	AA881330
21	75	52.1	158	5	AA807281
22	75	52.1	158	5	AA807275
23	75	52.1	693	3	AA92649
24	75	52.1	750	3	AA92645
25	75	52.1	750	3	AA92627

26	75	52.1	750	3	AA92640	AA92640	Mutant hu
27	75	52.1	750	3	AA92630	AA92630	Mutant hu
28	75	52.1	750	3	AA92646	AA92646	Mutant hu
29	75	52.1	750	3	AA92641	AA92641	Mutant hu
30	75	52.1	28	5	AAU11422	AAU11422	Synthetic
31	74	51.4	31	5	AA879180	AA879180	Human cyt
32	74	51.4	32	3	AA82636	AA82636	Tetanus t
33	74	51.4	126	3	AA845490	AA845490	Modified
34	74	51.4	137	3	AA82634	AA82634	Tetanus t
35	74	51.4	514	6	AA030491	AA030491	Human TNF
36	74	51.4	514	6	AA030490	AA030490	Human TNF
37	74	51.4	514	6	AA030495	AA030495	Human TNF
38	74	51.4	517	6	AA030492	AA030492	Human TNF
39	73	50.7	25	3	AA92650	AA92650	PSMpep007
40	73	50.7	31	2	AAW06129	AAW06129	Anti-cho
41	73	50.7	31	2	AAU02470	AAU02470	Fusion of
42	73	50.7	693	3	AA92647	AA92647	Mutant hu
43	73	50.7	750	3	AA92628	AA92628	Mutant hu
44	73	50.7	750	3	AA92644	AA92644	Mutant hu
45	72	50.0	17	2	AA62692	AA62692	Helper T

ALIGNMENTS

RESULT 1	AA884738	standard; peptide; 25 AA.
ID	AA884738	
AC	AA884738;	
DT	17-SEP-2001	(first entry)
DE	Amino acid sequence of lipopeptide M2-K.	
XX	Lipid-tailed protein; mucosal membrane; immune system; lipoprotein;	
KW	B cell response; T cell response; intranasal immunisation;	
KX	sublingual immunisation.	
XX		
OS	Synthetic.	
FH	Key	Location/Qualifiers
FT	Modified-site	1
FT		/note= "amidated residue with palmitoyl lipid residue
FT		with 16 carbon atoms attached via an epsilon bond"
XX		
PN	MO200141797-AZ.	
XX		
PD	14-JUN-2001.	
XX		
PF	08-DEC-2000; 2000WO-EP012794.	
XX		
PR	09-DEC-1999; 99US-0169952P.	
XX		
PA	(INSP) INST PASTEUR.	
PA	(UYLI-) UNIV LILLE 2.	
XX		
PI	Drulhe P, Gras-Masse H, Benmohamed L;	
XX		
DR	WPI; 2001-451597/48.	
XX		
PT	Inducing systemic immune response comprises mucosal administration of	
XX	lipid-tailed proteins or lipopeptide in absence of adjuvants.	
PS	Claim 22; Page 26; 34pp; English.	
XX		
CC	The present sequence represents a lipid-tailed protein. The lipid moiety	
CC	is able to cross the mucosal membrane, and deliver an antigen to the	
CC	immune system. The lipoprotein is administered to a mucosal membrane to	
CC	induce an immune response. The lipoprotein induces a B cell and/or a T	
CC	cell response. The lipoprotein is used to provide intranasal or	
XX	sublingual immunization	

Seq Sequence 25 AA; 77.4%; Score 111.5; DB 4; Length 25;
Query Match Best Local Similarity 96.0%; Pred. No. 5.5e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GROYT-KANSKFIGITERGRILKEP 24
DB 1 GROYIKRANSKFIGITERGRILKEP 25
RESULT 2
AAVS3552
ID AAVS3552 standard; protein; 28 AA.
AC AAVS3552;
XX
XX 18-JAN-2000 (first entry)
DT
DE Lipopeptide #3.
XX
XX Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
KW electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
KM human immunodeficiency virus; hepatitis B virus; papilloma virus;
KM melanoma; malaria; parasite.
XX
OS Synthetic.
OS Homo sapiens.
FH
FH Key Location/Qualifiers
FT Modified-site 1 /note="contains palmitoyl residue attached at the N-terminus and on the epsilon carbon of the side chain"
FT
PN FR2776926-A1.
XX
XX 08-OCT-1999.
PD
XX 07-APR-1998; 98FR-00004323.
PF
XX 07-APR-1998; 98FR-00004323.
PR
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (CNRS) CNRS CENT NAT RECH SCI.
PA (INSP) INST PASTEUR LILLE.
PI
PI Le Gal FA, Guillet JG, Gahery SH, Gras MH, Melnyk O, Tartar A;
DR WPI; 1999-583113/50.
XX
XX New lipopeptide containing lipid regions and two epitopes, all separated
PT by peptide spacers that impart hydrophilicity, useful in vaccines.
XX
XX Example 1; Page 9; 35pp; French.
PS
XX The invention relates to the generation of a lipopeptide comprising at
CC least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)
CC epitope and at least one lipid residue with (i) the epitopes and lipid
CC portion and (ii) the epitopes, being separated independently by peptide
CC spacers. These spacers comprise sequences of amino acids which carry an
CC overall electrical charge in neutral media to ensure that the lipopeptide
CC is hydrophilic. This peptide represents an example of them lipopeptide of
CC the invention. It contains 2 lipid residues attached at the N-terminal
CC end of the molecule and 2 epitopes separated by the spacer residues Ser-
CC Ser and Ala-Ala-Ala. Peptides AAVS3301-Y53549 represent peptide epitopes
CC used in the generation of the lipopeptides. These are used in therapeutic
CC or prophylactic compositions and vaccines to induce specific immune
CC responses against human immunodeficiency, hepatitis B or papilloma
CC viruses, p53 of melanoma or the malaria parasite
XX
XX Sequence 28 AA;
SQ
Query Match 67.4%; Score 97; DB 2; Length 28;

Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GROYIKANSKFIGITERGR 19
DB 1 GROYIKANSKFIGITERGR 19
RESULT 3
AAVS3551
ID AAVS3551 standard; protein; 29 AA.
XX
XX AAVS3551;
AC
XX 18-JAN-2000 (first entry)
DT
DE Lipopeptide #2.
XX
XX Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
KW electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
KM human immunodeficiency virus; hepatitis B virus; papilloma virus;
KM melanoma; malaria; parasite.
XX
OS Synthetic.
OS Homo sapiens.
FH
FH Key Location/Qualifiers
FT Modified-site 1 /note="contains palmitoyl residue attached at the N-terminus and on the epsilon carbon of the side chain"
FT
PN FR2776926-A1.
XX
XX 08-OCT-1999.
PD
XX 07-APR-1998; 98FR-00004323.
PF
XX 07-APR-1998; 98FR-00004323.
PR
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (CNRS) CNRS CENT NAT RECH SCI.
PA (INSP) INST PASTEUR LILLE.
PI
PI Le Gal FA, Guillet JG, Gahery SH, Gras MH, Melnyk O, Tartar A;
DR WPI; 1999-583113/50.
XX
XX New lipopeptide containing lipid regions and two epitopes, all separated
PT by peptide spacers that impart hydrophilicity, useful in vaccines.
XX
XX Example 1; Page 9; 35pp; French.
PS
XX The invention relates to the generation of a lipopeptide comprising at
CC least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)
CC epitope and at least one lipid residue with (i) the epitopes and lipid
CC portion and (ii) the epitopes, being separated independently by peptide
CC spacers. These spacers comprise sequences of amino acids which carry an
CC overall electrical charge in neutral media to ensure that the lipopeptide
CC is hydrophilic. This peptide represents an example of them lipopeptide of
CC the invention. It contains 2 lipid residues attached at the N-terminal
CC end of the molecule and 2 epitopes separated by the spacer residues Ser-
CC Ser and Ala-Ala-Ala. Peptides AAVS3301-Y53549 represent peptide epitopes
CC used in the generation of the lipopeptides. These are used in therapeutic
CC or prophylactic compositions and vaccines to induce specific immune
CC responses against human immunodeficiency, hepatitis B or papilloma
CC viruses, p53 of melanoma or the malaria parasite
XX
XX Sequence 29 AA;
SQ
Query Match 59.7%; Score 86; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 3 OYKANSKFIGITERGR 19
DB 4 OYKANSKFIGITERGR 20

RESULT 4
ABB79182
ID ABB79182 standard; peptide; 29 AA.
AC ABB79182;
XX
XX 07-AUG-2002 (first entry)
XX
DE Human cytomegalovirus PPI50 related vaccine peptide SEQ ID NO:8.
XX
XX Human cytomegalovirus PPI50; HCMV; human leukocyte antigen; HLA; CTL;
XX cytotoxic T lymphocyte; immunoreactive; epitope; vaccine; infection;
XX immunogenic; virucide; immune response.
XX
OS Human cytomegalovirus.
XX
XX WO200234769-A2.
XX
XX 02-MAY-2002.
XX
XX 22-OCT-2001; 2001MO-US032589.
XX
XX 20-OCT-2000; 2000US-0241944P.
XX
XX (CITY ) CITY OF HOPE.
XX
XX Diamond DJ;
XX
XX WPI; 2002-471432/50.
XX
XX New immunoreactive peptide cytotoxic T-lymphocytes epitope of human
XX cytomegalovirus PPI50, useful for creating vaccines against
XX cytomegalovirus infection.
XX
XX Disclosure; Page 14; 28pp; English.
XX
XX The present invention describes a peptide (I) which is an immunogenic
XX epitope recognised by CD8+ class I major histocompatibility complex (MHC)
XX restricted cytotoxic T-lymphocytes (CTL) of patients harbouring latent
XX cytomegalovirus (HCMV) infection. Also described are: (1) a vaccine (II)
XX against human cytomegalovirus (CMV), which comprises (I); (2) a cellular
XX vaccine (III) against human CMV comprises antigen presenting cells which
XX present (I); (3) a recombinant viral vector (IV) which contains a gene
XX encoding (I); and (4) an immunological reagent (V) which comprises (I).
XX (I) has virucide activity. (II) or (III) are useful for modulating the
XX immune response to HCMV infection, and for vaccinating a mammal in need
XX of it against HCMV. (I) is useful for eliciting a cellular immune against
XX HCMV by normal and immunodeficient subjects, for creating efficient
XX vaccines against HCMV, and in immunological methods to detect PPI50-
XX reactive CTL in a patient or a sample from the patient. The present
XX sequence represents a HCMV PPI50 related vaccine peptide, which is given
XX in the exemplification of the present invention
XX
XX Sequence 29 AA;
XX
XX Query Match 56.2%; Score 81; DB 5; Length 29;
XX Best Local Similarity 100.0%; Pred. No. 7e-06;
XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
AC AAY92665;
XX
XX 10-AUG-2000 (first entry)
XX
XX MUC-1 analogue containing foreign epitopes.
XX
XX Mucin repeat; MUC-1 analogue; vaccination; self-protein; cancer;
XX cytotoxic T-lymphocyte immunity; breast cancer; prostate cancer;
XX cell-associated peptide antigen; foreign epitope.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 61..75
XX /label= P2
XX Peptide 136..156
XX /label= P30
XX /note= "q"
XX
XX WO20020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK000525.
XX
XX 05-OCT-1998; 98DK-00001261.
XX
XX 20-OCT-1998; 98US-0105011P.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated peptide
XX antigens for the treatment of breast and prostate cancer.
XX
XX Example 4; Page; 220pp; English.
XX
XX This is an immunogenized MUC-1 analogue containing foreign epitopes P2
XX and P30. Immunogenic analogues of MUC-1 and, e.g. human prostate specific
XX membrane antigen (hPSM) can be used in the claimed method as an
XX autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
XX binding regions and cysteine residues involved in disulfide bonds are
XX preserved in the immunogenized forms (see features table). 10 regions
XX suitable for the insertion of foreign T helper epitopes were identified.
XX The method is used for inducing immune responses against weakly
XX immunogenic cell-associated peptide antigens (PA) such as those
XX associated with cancers (self-proteins), e.g. hPSM, heregulin 2 (Her2)
XX and/or fibroblast growth factor 8b (FGF8b). The method comprises
XX effecting simultaneous presentation by antigen producing cells (APCs) of
XX the animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
XX group derived from the PA and/or at least 1 B-cell group derived from the
XX cell-associated PA; and (2) at least 1 first T helper cell group which is
XX foreign to the animal. Analogues of human PSM, human Her2 and
XX human/murine FGF8b comprising a substantial part of all known and
XX predicted CTL and B-cell epitopes of the respective PA and including at
XX least one foreign T helper epitope are also claimed. The method is used
XX to treat prostate, prostate/breast or breast cancer when the PA is human
XX PSM, FGF8b and Her2, respectively. Note: This sequence does not appear in
XX the specification. It was made using the mucin repeat sequence
XX (AAY92664), P2 and P30 (AAY92625-26), which appear on pages 220, 213 and
XX 214 respectively, of the specification
XX
XX Sequence 216 AA;
XX
XX Query Match 54.9%; Score 79; DB 3; Length 216;
XX Best Local Similarity 65.5%; Pred. No. 0.00016;
XX Matches 19; Conservative 0; Mismatches 8; Indels 2; Gaps 1;
```

Db 60 ROYIKANSKFIGITELPAPGSTADPAHGV 88

RESULT 6
ID AAY82632
XX AAY82632 standard; peptide; 31 AA.
XX AAY82632;
DE 07-AUG-2000 (first entry)
XX
XX Tetanus toxoid T cell epitope and Der pii B cell epitope peptide.
XX
XX T cell epitope; B cell epitope; allergy; allergen; antigenic;
XX anti-allergic; antiallergic; antiinflammatory; dermatological;
XX immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
XX atopic dermatitis; acute urticaria; chronic urticaria;
XX gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
XX anaphylactic reaction; drug hypersensitivity; allergic reaction.
XX
XX Dermatophagoides pteromyssinus.
OS Clostridium tetani.
OS Synthetic.
XX
XX WO200006694-A2.
XX
XX 10-FEB-2000.
XX
XX 20-JUL-1999; 99WO-BE000092.
XX
XX 30-JUL-1998; 98EP-00870167.
XX
XX (UNIO) DCB SA.
XX
XX Saint-Remy J, Jacquemin M;
XX
XX WPI; 2000-422470/36.
XX
XX New compound for prevention and treatment of allergies comprises at least
XX one allergen antigenic determinant recognized by a B cell and at least
XX one antigenic determinant which does not trigger T cell activation.
XX
XX Claim 8; Page 35; 50pp; English.

The present invention describes a compound (I) for the prevention and/or treatment of allergy. The compound comprises at least one allergen antigenic determinant (i) recognised by a B cell or an antibody secreted by a B cell of a non-atopic individual and at least one antigenic determinant (ii) different from the allergen that triggers T cell activation. (I) has antiallergic, antiasthmatic, antiinflammatory, dermatological and immunosuppressive activities, and can be used in a vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to treat and/or prevent allergies or a disease of allergic origin, especially hypersensitivities. These include rhinitis, sinusitis, bronchial asthma, atopic dermatitis, some forms of acute and chronic urticaria, gastro-intestinal syndromes associated with the ingestion of food allergens, oro-pharyngeal syndrome, anaphylactic reactions associated with drug hypersensitivities and/or a mixture of these. The use of (I) in the treatment of allergic conditions avoids the need for drug treatment, which often causes undesirable side-effects. Also, prior art drug therapies alleviate symptoms, but do not influence their causes, however (I) actually combats the cause of an allergic reaction. The present sequence represents a specifically claimed compound peptide sequence from the present invention

Sequence 31 AA;
54.5%; Score 78.5; DB 3; Length 31;
Best Local Similarity 60.0%; Pred. No. 2e-05;
Matches 18; Conservative 2; Mismatches 5; Indels 5; Gaps 1

3 OYIKANSKFIGITTEG-----RIKKEPVHG 27
|||||KANSKFIGITTEG-----RIKKEPVHG 27
:::||||

Db 1 QYIKANSKFIGITELGHEIRKKVLPQCHG 30

RESULT 7
AA92639
ID AA92639 standard; protein; 750 AA.
XX
XX AA92639;
AC
DT 10-AUG-2000 (first entry)
DE Mutant human prostate specific membrane antigen construct, hPSM5.1.
XX
XX Prostate specific membrane antigen; immunogenized construct; mutant;
XX vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KM prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
XX Homo sapiens.
OS Synthetic.
CS
FH Key Location/Qualifiers
FT 21..41
FT /label= P30
FT /note= "foreign epitope"
FT 305..319
FT Peptide /label= P2
FT /note= "foreign epitope"
FT
FN WO200020027-A2.
PN
PD 13-APR-2000.
PP 05-OCT-1999; 99WO-DK000525.
PR 05-OCT-1998; 98DK-00001261.
PR 20-OCT-1998; 98US-0105011P.
PX (MEBI-) M & B BIOTECH AS.
PA
PI Steinaa I., Mouritsen S., Nielsen KG., Haaning J., Leach D., Dalum I.;
PI Gautam A., Birke P., Karlsson G;
DR WPI, 2000-349917/30.
PT Inducing immune responses to weakly immunogenic, tumor associated peptide
PS antigens for the treatment of breast and prostate cancer.
XX
XX Example 1; Page; 220pp; English.

AA92639 are mutant immunogenized human prostate specific membrane antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30). The immunogenic analogues of PSM can be used in the claimed method as an autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody binding regions and cysteine residues involved in disulfide bonds are preserved in the immunogenized forms. The method is used for inducing immune responses against weakly immunogenic cell-associated peptide antigens (PA) such as those associated with cancers (self-proteins), e.g./or human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The method comprises effecting simultaneous presentation by antigen producing cells (APCs) of the animal's immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1 B-cell group derived from the cell-associated PA; and (2) at least 1 first T helper cell group which is foreign to the animal. Analogues of human PSM, human Her2 and human/murine FGF8b comprising a substantial part of all known and predicted CTL and B-cell epitopes of the respective PA and including at least one foreign T helper epitope are also claimed. The method is used to treat prostate, prostate/breast or breast cancer when the PA is human PSM, FGF8b and Her2, respectively. Note: This sequence was constructed from the wild type human PSM (AA92639), which appears on pages 184-187 of the specification

SQ Sequence 750 AA;


```

XX Universal immunostimulator having GG spacers.
XX Helper T cell epitope; universal immune stimulator; invasin; hapten;
XX tetanus toxin.
XX Synthetic.
XX Key Location/Qualifiers
XX FT 1..16
XX FT Domain /note= "invasin domain"
XX FT 19..35
XX FT Domain /note= "tetanus toxin helper T cell epitope"
XX WO9425060-A1.
XX PD 10-NOV-1994.
XX PF 28-APR-1994; 94WO-US004832.
XX PR 27-APR-1993; 93US-00057166.
XX PR 14-APR-1994; 94US-00229275.
XX PA (LADD/) LADD A E.
XX PA (WANG/) WANG C Y.
XX PA (ZAMB/) ZAMB T.
XX PI Ladd AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
XX PT Immunogenic luteinising hormone releasing hormone peptide(s) - that
XX PT suppress LHRH activity in males and females.
XX PS Disclosure; Page 95; 213pp; English.
XX CC Synthetic immunogenic peptides are provided in which a universal immune
XX CC stimulator is linked to a peptide or protein hapten containing B cell
XX CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX CC potent immune responses to the coupled peptide or protein. The stimulator
XX CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
XX CC immune response to the coupled peptide in members of a heterogeneous
XX CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
XX CC sequence from the invasin protein of Yersinia. Spacer amino acid
XX CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
XX CC domains and between the immune stimulator and hapten components. When the
XX CC hapten is LHRH, then optionally the invasin domain can be omitted from
XX CC the immune stimulator component. The present sequence is an example of an
XX CC invasin-GG-Th-GG- immune stimulator to which a hapten can be bonded.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 37 AA;
XX
XX Query Match 52.8%; Score 76; DB 2; Length 37;
XX Best Local Similarity 88.2%; Pred. No. 6.2e-05;
XX Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 RQYIKANSKFIQTTERG 18
XX :|||||
XX Db 20 KQYIKANSKFIQTTELG 36
XX
XX RESULT 11
XX ID AAR62723 standard; peptide; 47 AA.
XX AC AAR62723;
XX DT 25-MAR-2003 (revised)
XX DT 17-SEP-1995 (first entry)
XX LHRH-containing immunogenic peptide.
XX

```

```

XX Helper T cell epitope; universal immune stimulator; invasin; hapten;
XX vaccine; LHRH; luteinising hormone releasing hormone; prostate;
XX androgen-dependent carcinoma; antitumour; infertility; tetanus toxin.
XX Synthetic.
XX Key Location/Qualifiers
XX FT 1..16
XX FT Domain /note= "invasin domain"
XX FT 19..35
XX FT Domain /note= "tetanus toxin helper T cell epitope"
XX FT 38..47
XX FT Domain /note= "LHRH hapten"
XX WO9425060-A1.
XX PD 10-NOV-1994.
XX PF 28-APR-1994; 94WO-US004832.
XX PR 27-APR-1993; 93US-00057166.
XX PR 14-APR-1994; 94US-00229275.
XX PA (LADD/) LADD A E.
XX PA (WANG/) WANG C Y.
XX PA (ZAMB/) ZAMB T.
XX PI Ladd AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
XX PT Immunogenic luteinising hormone releasing hormone peptide(s) - that
XX PT suppress LHRH activity in males and females.
XX PS Claim 8; Page 88; 213pp; English.
XX CC Synthetic immunogenic peptides are provided in which a universal immune
XX CC stimulator is linked to a peptide or protein hapten containing B cell
XX CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX CC potent immune responses to the coupled peptide or protein. The stimulator
XX CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
XX CC immune response to the coupled peptide in members of a heterogeneous
XX CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
XX CC sequence from the invasin protein of Yersinia. Spacer amino acid
XX CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
XX CC domains and between the immune stimulator and hapten components. When the
XX CC hapten is LHRH, then optionally the invasin domain can be omitted from
XX CC the immune stimulator component. The present sequence represents an LHRH-
XX CC containing immunogenic peptide as above which can be used as a potent
XX CC vaccine for treating e.g. prostatic hyperplasia, androgen-dependent
XX CC carcinoma, prostatic carcinoma, testicular carcinoma, androgen-dependent
XX CC benign uterine tumours, recurrent functional ovarian cysts, (severe)
XX CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
XX CC induction of infertility. (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 47 AA;
XX
XX Query Match 52.8%; Score 76; DB 2; Length 47;
XX Best Local Similarity 88.2%; Pred. No. 8.2e-05;
XX Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 RQYIKANSKFIQTTERG 18
XX :|||||
XX Db 20 KQYIKANSKFIQTTELG 36
XX
XX RESULT 12
XX ID AAY92638 standard; protein; 750 AA.
XX AC AAY92638;
XX DT 10-AUG-2000 (first entry)
XX

```

```

XX DE Mutant human prostate specific membrane antigen construct, hpsm3.1.
XX XX
XX KW Prostate specific membrane antigen; immunogenized construct; mutant;
XX KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
XX KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX FH 21..41
XX FT /label= P30
XX FT /note="Foreign epitope"
XX FT 213..227
XX FT /label= P2
XX FT /note="foreign epitope"
XX
XX WO200020027-A2.
XX
XX PD 13-APR-2000.
XX
XX PF 05-OCT-1999; 99WO-DK000525.
XX
XX PR 05-OCT-1998; 98DK-00001261.
XX PR 20-OCT-1998; 98US-0105011P.
XX
XX PA (MEBI-) M & E BIOTECH AS.
XX
XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX PI Gautam A, Birk P, Karlsson G,
XX
XX DR WPI; 2000-349917/30.
XX
XX PT Inducing immune responses to weakly immunogenic, tumor associated peptide
XX PT antigens for the treatment of breast and prostate cancer.
XX
XX PS Example 1; Page; 220pp; English.
XX
XX CC AAY92627-49 are mutant immunogenized human prostate specific membrane
XX CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
XX CC The immunogenic analogues of PSM can be used in the claimed method as an
XX CC autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
XX CC binding regions and cysteine residues involved in disulfide bonds are
XX CC preserved in the immunogenized forms. The method is used for inducing
XX CC immune responses against weakly immunogenic cell-associated peptide
XX CC antigens (PA) such as those associated with cancers (self-proteins), e.g.
XX CC human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
XX CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
XX CC simultaneous presentation by antigen producing cells (APCs) of the
XX CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
XX CC group derived from the PA and/or at least 1 B-cell group derived from the
XX CC cell-associated PA; and (2) at least 1 first T helper cell group which is
XX CC foreign to the animal. Analogues of human PSM, human Her2 and
XX CC human/murine FGF8b comprising a substantial part of all known and
XX CC predicted CTL and B-cell epitopes of the respective PA and including at
XX CC least one foreign T helper epitope are also claimed. The method is used
XX CC to treat prostate, prostate/breast or breast cancer when the PA is human
XX CC PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
XX CC from the wild type human PSM (AAY92619), which appears on pages 184-187
XX CC of the specification
XX
XX SQ Sequence 750 AA;

```

```

RESULT 13
AAY92652
ID AAY92652 standard; peptide; 25 AA.
XX
XX AC AAY92652;
XX
XX DT 10-AUG-2000 (first entry)
XX
XX DE PSMpep009 - P2 inserted in hpsm insertion position 10.
XX
XX KW Foreign epitope; P2; prostate specific membrane antigen; vaccination;
XX KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
XX KW prostate cancer; cell-associated peptide antigen.
XX
XX OS Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH 6..20
XX FT /label= P2
XX FT
XX
XX WO200020027-A2.
XX
XX PD 13-APR-2000.
XX
XX PF 05-OCT-1999; 99WO-DK000525.
XX
XX PR 05-OCT-1998; 98DK-00001261.
XX PR 20-OCT-1998; 98US-0105011P.
XX
XX PA (MEBI-) M & E BIOTECH AS.
XX
XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX PI Gautam A, Birk P, Karlsson G,
XX
XX DR WPI; 2000-349917/30.
XX
XX PT Inducing immune responses to weakly immunogenic, tumor associated peptide
XX PT antigens for the treatment of breast and prostate cancer.
XX
XX PS Example 1; Page 117; 220pp; English.
XX
XX CC AAY92650-55 are peptides designed which correspond to the P2 and P30
XX CC epitopes with 5 flanking human prostate specific membrane antigen (hpsm)
XX CC amino acids in each end. The flanking amino acids correspond to the
XX CC epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g. T
XX CC cell proliferation assays, but also for ELISA or other in vitro assays.
XX CC The claims detail a method for inducing immune responses against weakly
XX CC immunogenic cell-associated peptide antigens (PA) such as those
XX CC associated with cancers (i.e. self-proteins), for example, hpsm,
XX CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The method
XX CC comprises effecting simultaneous presentation by antigen producing cells
XX CC (APCs) of the animals immune system of: (1) at least 1 CTL (cytotoxic T-
XX CC lymphocyte) group derived from the PA and/or at least 1 B-cell group
XX CC derived from the cell-associated PA; and (2) at least 1 first T helper
XX CC cell group which is foreign to the animal. Analogues of human PSM, human
XX CC Her2 and human/murine FGF8b comprising a substantial part of all known
XX CC and predicted CTL and B-cell epitopes of the respective PA and including
XX CC at least one foreign T helper epitope (e.g. P2 and/or P30) are also
XX CC claimed. The method is used to treat prostate, prostate/breast or breast
XX CC cancer when the PA is human PSM, FGF8b and Her2, respectively
XX
XX SQ Sequence 25 AA;

```

```

Query Match 52.8%; Score 76; DB 3; Length 750;
Best Local Similarity 93.8%; Pred. No. 0.0021;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Query Match 52.1%; Score 75; DB 3; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 14
 ID AAY92651 standard; peptide; 25 AA.
 AC AAY92651;
 XX 10-AUG-2000 (first entry)
 DT
 DE PSM008 - P2 inserted in hPSM insertion position 8.
 XX
 XX Foreign epitope; P2; prostate specific membrane antigen; vaccination;
 KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
 KW prostate cancer; cell-associated peptide antigen.
 XX
 XX Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT peptide 6.20
 FT /label= P2
 WO200020027-A2.
 XX 13-APR-2000.
 XX
 XX 05-OCT-1999; 99MO-DK000525.
 XX
 XX 05-OCT-1998; 98DK-00001261.
 XX 20-OCT-1998; 98US-0105011P.
 XX
 XX (MEBI-) M & E BIOTECH AS.
 PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 PI
 DR WPI; 2000-349917/30.
 XX
 PT Inducing immune responses to weakly immunogenic, tumor associated peptide
 PT antigens for the treatment of breast and prostate cancer.
 XX
 XX Example 1; Page 117; 220pp; English.
 XX
 CC AAY92650-55 are peptides designed which correspond to the P2 and P30
 CC epitopes with 5 flanking human prostate specific membrane antigen (hPSM)
 CC amino acids in each end. The flanking amino acids correspond to the
 CC epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g. T
 CC cell proliferation assays, but also for ELISA or other in vitro assays.
 CC The claims detail a method for inducing immune responses against weakly
 CC immunogenic cell-associated peptide antigens (PA) such as those
 CC associated with cancers (i.e. self-proteins), for example, hPSM,
 CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The method
 CC comprises effecting simultaneous presentation by antigen producing cells
 CC (APCs) of the animals immune system of: (1) at least 1 CTL (cytotoxic T-
 CC lymphocyte) group derived from the PA and/or at least 1 B-cell group
 CC derived from the cell-associated PA; and (2) at least 1 first T helper
 CC cell group which is foreign to the animal. Analogues of human PSM, human
 CC Her2 and human/murine FGF8b comprising a substantial part of all known
 CC and predicted CTL and B-cell epitopes (e.g. P2 and/or P30) are also
 CC claimed. The method is used to treat prostate, prostate/breast or breast
 CC cancer when the PA is human PSM, FGF8b and Her2, respectively
 XX
 XX Sequence 25 AA:
 SQ

Query Match 52.1%; Score 75; DB 3; Length 25;
 Best Local Similarity 100.0%; Pred. No. 5.7e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYIKANSKFIGITTE 16
 |||||
 DB 5 RYIKANSKFIGITTE 19

RESULT 15
 ID ABB79179 standard; peptide; 31 AA.
 AC ABB79179;
 XX 07-AUG-2002 (first entry)
 DT
 DE Human cytomegalovirus PP150 related vaccine peptide SEQ ID NO:5.
 XX
 XX Human cytomegalovirus PP150; HCMV; human leukocyte antigen; HLA; CTL;
 KW cytotoxic T lymphocyte; immunoreactive; epitope; vaccine; infection;
 KW immunogenic; virucide; immune response.
 XX
 XX Human cytomegalovirus.
 OS
 XX
 XX WO200234769-A2.
 XX
 XX 02-MAY-2002.
 XX
 XX 22-OCT-2001; 2001WO-US032589.
 XX
 XX 20-OCT-2000; 2000US-0241944P.
 XX
 XX (CITY) CITY OF HOPE.
 PA
 PI Diamond DJ;
 PI
 DR WPI; 2002-471432/50.
 XX
 PT New immunoreactive peptide cytotoxic T-lymphocytes epitope of human
 PT cytomegalovirus PP150, useful for creating vaccines against
 PT cytomegalovirus infection.
 XX
 XX Disclosure; Page 14; 28pp; English.
 XX
 CC The present invention describes a peptide (I) which is an immunogenic
 CC epitope recognised by CD8+ class I major histocompatibility complex (MHC)
 CC restricted cytotoxic T-lymphocytes (CTL) of patients harbouring latent
 CC cytomegalovirus (HCMV) infection. Also described are: (1) a vaccine (II)
 CC against human cytomegalovirus (CMV), which comprises (i); (2) a cellular
 CC vaccine (III) against human CMV comprises antigen presenting cells which
 CC present (i); (3) a recombinant viral vector (IV) which contains a gene
 CC encoding (i); and (4) an immunological reagent (V) which comprises (i).
 CC (i) has virucide activity. (ii) or (iii) are useful for modulating the
 CC immune response to HCMV infection, and for vaccinating a mammal in need
 CC of it against HCMV. (I) is useful for eliciting a cellular immune against
 CC HCMV by normal and immunodeficient subjects, for creating efficient
 CC vaccines against HCMV, and in immunological methods to detect PP150-
 CC reactive CTL in a patient or a sample from the patient. The present
 CC sequence represents a HCMV PP150 related vaccine peptide, which is given
 CC in the exemplification of the present invention
 XX
 XX Sequence 31 AA:
 SQ

Query Match 52.1%; Score 75; DB 5; Length 31;
 Best Local Similarity 63.0%; Pred. No. 7.4e-05;
 Matches 17; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 3 QYIKANSKFIGITERG-RILKEPVHG 27
 |||||
 DB 4 QYIKANSKFIGITEAAAGTSTPVOG 30

Search completed: July 20, 2004, 06:30:27
 Job time : 42.7705 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:10:18 ; Search time 11.0164 Seconds

(without alignments)
279.413 Million cell updates/sec

Title: US-09-673-166B-275

Sequence: 168
1 GROYTKANSKFLGITRGRNPDIVYQYMDL 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query length	DB ID	Description
1	78	46.4	210	2 F47330	RNA-directed DNA p
2	75	44.6	210	2 D47330	RNA-directed DNA p
3	75	44.6	210	2 A47330	RNA-directed DNA p
4	75	44.6	210	2 B47330	RNA-directed DNA p
5	75	44.6	210	2 C47330	RNA-directed DNA p
6	75	44.6	210	2 E47330	RNA-directed DNA p
7	75	44.6	219	2 S32139	RNA-directed DNA p
8	75	44.6	219	2 S32071	RNA-directed DNA p
9	75	44.6	219	2 S32079	RNA-directed DNA p
10	75	44.6	219	2 S32077	RNA-directed DNA p
11	75	44.6	219	2 S32075	RNA-directed DNA p
12	75	44.6	219	2 S32073	RNA-directed DNA p
13	75	44.6	219	2 S32070	RNA-directed DNA p
14	75	44.6	219	2 S32072	RNA-directed DNA p
15	75	44.6	219	2 S32066	RNA-directed DNA p
16	75	44.6	219	2 S32064	RNA-directed DNA p
17	75	44.6	219	2 S32062	RNA-directed DNA p
18	75	44.6	219	2 S32060	RNA-directed DNA p
19	75	44.6	219	2 S32058	RNA-directed DNA p
20	75	44.6	219	2 S32056	RNA-directed DNA p
21	75	44.6	219	2 S32054	RNA-directed DNA p
22	75	44.6	219	2 S32052	RNA-directed DNA p
23	75	44.6	219	2 S32050	RNA-directed DNA p
24	75	44.6	219	2 S32048	RNA-directed DNA p
25	75	44.6	219	2 S32046	RNA-directed DNA p
26	75	44.6	219	2 S32044	RNA-directed DNA p
27	75	44.6	219	2 S32042	RNA-directed DNA p
28	75	44.6	219	2 S32040	RNA-directed DNA p
29	75	44.6	219	2 S32038	RNA-directed DNA p

30	75	44.6	219	2 S32095	RNA-directed DNA p
31	75	44.6	219	2 S32056	RNA-directed DNA p
32	75	44.6	219	2 S32064	RNA-directed DNA p
33	75	44.6	219	2 S32054	RNA-directed DNA p
34	75	44.6	219	2 S32062	RNA-directed DNA p
35	75	44.6	219	2 S32061	RNA-directed DNA p
36	75	44.6	219	2 S32058	RNA-directed DNA p
37	75	44.6	219	2 S32060	RNA-directed DNA p
38	75	44.6	219	2 S32057	RNA-directed DNA p
39	75	44.6	219	2 S32159	RNA-directed DNA p
40	75	44.6	219	2 S32157	RNA-directed DNA p
41	75	44.6	219	2 S32160	RNA-directed DNA p
42	75	44.6	219	2 S32078	RNA-directed DNA p
43	75	44.6	219	2 S32135	RNA-directed DNA p
44	75	44.6	219	2 S32069	RNA-directed DNA p
45	75	44.6	219	2 S32126	RNA-directed DNA p

ALIGNMENTS

RESULT 1

F47330
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (strain N1/Alternate names: DNA nucleotidyltransferase (RNA-directed); reverse transcriptase; rev C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 22-Apr-1995
C:Accession: F47330
R:Shirasaka, T.; Yarchuan, R.; O'Brien, M.C.; Husson, R.N.; Anderson, B.D.; Kojima, E.; Proc. Natl. Acad. Sci. U.S.A. 90, 562-566, 1993
A:Title: Changes in drug sensitivity of human immunodeficiency virus type 1 during therapy
A:Reference number: A47330; MUID:93133828; PMID:8380641
A:Accession: F47330
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-210 <SH1>
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match
Best Local Similarity 46.4%; Score 78; DB 2; Length 210;
Matches 18; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

Qy 9 SKFIGITE--RGRNPDIVYQYMDL 32
Db 142 SSMIKLPRKQNPDIYQYMDL 167

RESULT 2

D47330
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (strain N1/Alternate names: DNA nucleotidyltransferase (RNA-directed); reverse transcriptase; rev C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 22-Apr-1995
C:Accession: D47330
R:Shirasaka, T.; Yarchuan, R.; O'Brien, M.C.; Husson, R.N.; Anderson, B.D.; Kojima, E.; Proc. Natl. Acad. Sci. U.S.A. 90, 562-566, 1993
A:Title: Changes in drug sensitivity of human immunodeficiency virus type 1 during therapy
A:Reference number: A47330; MUID:93133828; PMID:8380641
A:Accession: D47330
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-210 <SH1>
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase; polymorphism

Query Match
Best Local Similarity 44.6%; Score 75; DB 2; Length 210;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 17 RGRNPDIVYQYMDL 32
Db 152 RKQNPDIYQYMDL 167

```

RESULT 3
A47330
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (strain
N/Alternate names: DNA nucleotidyltransferase (RNA-directed); reverse transcriptase; rev
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 22-Apr-1995
C/Accession: A47330
R/Shirasaka, T.; Yarchan, R.; O'Brien, M.C.; Husson, R.N.; Anderson, B.D.; Kojima, E.;
Proc. Natl. Acad. Sci. U.S.A. 90, 562-566, 1993
A/Title: Changes in drug sensitivity of human immunodeficiency virus type 1 during thera
A/Reference number: A47330; PMID:93133828; PMID:8380641
A/Accession: A47330
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-210 <SHI>
C/Superfamily: pol polyprotein
C/Keywords: nucleotidyltransferase

Query Match          44.6%; Score 75; DB 2; Length 210;
Best Local Similarity 87.5%; Pred. No. 0.0045;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
|:|||||
Db 152 RKNPDIIVYQYMDL 167

RESULT 4
B47330
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (strain
N/Alternate names: DNA nucleotidyltransferase (RNA-directed); reverse transcriptase; rev
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 22-Apr-1995
C/Accession: B47330
R/Shirasaka, T.; Yarchan, R.; O'Brien, M.C.; Husson, R.N.; Anderson, B.D.; Kojima, E.;
Proc. Natl. Acad. Sci. U.S.A. 90, 562-566, 1993
A/Title: Changes in drug sensitivity of human immunodeficiency virus type 1 during thera
A/Reference number: A47330; PMID:93133828; PMID:8380641
A/Accession: B47330
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-210 <SHI>
C/Superfamily: pol polyprotein
C/Keywords: nucleotidyltransferase; polymorphism

Query Match          44.6%; Score 75; DB 2; Length 210;
Best Local Similarity 87.5%; Pred. No. 0.0045;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
|:|||||
Db 152 RKNPDIIVYQYMDL 167

RESULT 5
C47330
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (strain
N/Alternate names: DNA nucleotidyltransferase (RNA-directed); reverse transcriptase; rev
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 22-Apr-1995
C/Accession: C47330
R/Shirasaka, T.; Yarchan, R.; O'Brien, M.C.; Husson, R.N.; Anderson, B.D.; Kojima, E.;
Proc. Natl. Acad. Sci. U.S.A. 90, 562-566, 1993
A/Title: Changes in drug sensitivity of human immunodeficiency virus type 1 during thera
A/Reference number: A47330; PMID:93133828; PMID:8380641
A/Accession: C47330
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-210 <SHI>
C/Superfamily: pol polyprotein
C/Keywords: nucleotidyltransferase; polymorphism

Query Match          44.6%; Score 75; DB 2; Length 210;
Best Local Similarity 87.5%; Pred. No. 0.0045;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
|:|||||
Db 152 RKNPDIIVYQYMDL 167

```

```

Query Match          44.6%; Score 75; DB 2; Length 210;
Best Local Similarity 87.5%; Pred. No. 0.0045;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
|:|||||
Db 152 RKNPDIIVYQYMDL 167

RESULT 6
B47330
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (strain
N/Alternate names: DNA nucleotidyltransferase (RNA-directed); reverse transcriptase; rev
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 22-Apr-1995
C/Accession: B47330
R/Shirasaka, T.; Yarchan, R.; O'Brien, M.C.; Husson, R.N.; Anderson, B.D.; Kojima, E.;
Proc. Natl. Acad. Sci. U.S.A. 90, 562-566, 1993
A/Title: Changes in drug sensitivity of human immunodeficiency virus type 1 during therai
A/Reference number: A47330; PMID:93133828; PMID:8380641
A/Accession: B47330
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-210 <SHI>
C/Superfamily: pol polyprotein
C/Keywords: nucleotidyltransferase; polymorphism

Query Match          44.6%; Score 75; DB 2; Length 210;
Best Local Similarity 87.5%; Pred. No. 0.0045;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
|:|||||
Db 152 RKNPDIIVYQYMDL 167

RESULT 7
S32139
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C/Accession: S32139
R/Wildemann, B.; Haas, J.; Hahn, M.; Ehrhart, K.; Storch-Hagenlocher, B.
submitted to the EMBL Data Library, February 1993
A/Description: In vivo occurrence of drug resistance mutations under prolonged zidovudine
A/Reference number: S32117
A/Accession: S32139
A/Molecule type: DNA
A/Residues: 1-219 <NTL>
A/Cross-references: EMBL:X71106; NID:G287982; PIDN:CA50423.1; PID:G287983
C/Superfamily: pol polyprotein
C/Keywords: nucleotidyltransferase

Query Match          44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
|:|||||
Db 172 RKNPDIIVYQYMDL 187

RESULT 8
S32071
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment
N/Alternate names: reverse transcriptase
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C/Accession: S32071
R/Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
submitted to the EMBL Data Library, January 1993
A/Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 infe

```

A:Reference number: S32047
A:Accession: S32071
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70606; NID:g287694; PIDN:CAA49960.1; PID:g938201
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYIYQYMDL 32
|:|||||||||
Db 172 RKQNPDIYIYQYMDL 187

RESULT 9
S32079
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S32079
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32079
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70610; NID:g287702; PIDN:CAA49964.1; PID:g938209
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYIYQYMDL 32
|:|||||||||
Db 172 RKQNPDIYIYQYMDL 187

RESULT 10
S32077
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S32077
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32077
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70609; NID:g287700; PIDN:CAA49963.1; PID:g938207
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYIYQYMDL 32
|:|||||||||
Db 172 RKQNPDIYIYQYMDL 187

RESULT 11
S32075

RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S32075
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32075
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70608; NID:g287698; PIDN:CAA49962.1; PID:g938205
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYIYQYMDL 32
|:|||||||||
Db 172 RKQNPDIYIYQYMDL 187

RESULT 12
S32074
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S32074
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32074
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70600; NID:g287697; PIDN:CAA49954.1; PID:g938204
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYIYQYMDL 32
|:|||||||||
Db 172 RKQNPDIYIYQYMDL 187

RESULT 13
S32073
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S32073
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32073
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70607; NID:g287696; PIDN:CAA49961.1; PID:g938203
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIVIQYMDDL 32
| : |||||
Db 172 RKQNPDIIVIQYMDDL 187

RESULT 14

S32070
RNA-directed DNA polymerase (BC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
A:Submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32070
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70598; NID:G287693; PIDN:CAA49952.1; PID:G938200
A:Accession: S32067
A:Molecule type: DNA
A:Residues: 1-219 <WIL2>
A:Cross-references: EMBL:X70596; NID:G287689; PIDN:CAA49950.1; PID:G938197
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIVIQYMDDL 32
| : |||||
Db 172 RKQNPDIIVIQYMDDL 187

RESULT 15

S32080
RNA-directed DNA polymerase (BC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
A:Submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32080
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70611; NID:G287703; PIDN:CAA49965.1; PID:G938210
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIVIQYMDDL 32
| : |||||
Db 172 RKQNPDIIVIQYMDDL 187

Search completed: July 20, 2004, 06:34:10
Job time : 11.0164 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:17 ; Search time 6.29508 Seconds

(without alignments)
264.690 Million cell updates/sec

Title: US-09-673-166b-275

Perfect score: 168
Sequence: 1 GRQYIKANSKFGITGRGNPDIVITYQWDDL 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	44.6	1003	1	POL_HV1A2
2	75	44.6	1003	1	POL_HV1H2
3	75	44.6	1003	1	POL_HV1N5
4	75	44.6	1003	1	POL_HV1OY
5	75	44.6	1006	1	POL_HV1MN
6	75	44.6	1015	1	POL_HV1B5
7	75	44.6	1015	1	POL_HV1BR
8	75	44.6	1015	1	POL_HV1PV
9	74	44.0	1007	1	POL_HV1JR
10	73	43.5	1003	1	POL_HV1Y2
11	72	42.9	1002	1	POL_HV1Y2
12	72	42.9	1002	1	POL_HV1MA
13	72	42.9	1015	1	POL_HV1U4
14	72	42.9	1027	1	POL_HV1B1
15	71	42.3	1002	1	POL_SIVCZ
16	71	42.3	1002	1	POL_HV1ND
17	70	42.3	1002	1	POL_HV1Z2
18	70	42.3	1002	1	POL_HV1Z2
19	68	40.5	1002	1	TEYX_CLOTE
20	68	40.5	1002	1	POL_HV1EL
21	68	40.5	1002	1	POL_HV1RH
22	60	35.7	1035	1	POL_HV2SB
23	60	35.7	1035	1	POL_HV2ST
24	60	35.7	1035	1	POL_HV2ST
25	59	35.1	1035	1	POL_HV2ST
26	59	35.1	1035	1	POL_HV2ST
27	59	35.1	1035	1	POL_HV2ST
28	58	34.5	1145	1	POL_HV2BE
29	58	34.5	1145	1	POL_HV2BE
30	58	34.5	1145	1	POL_HV2BE
31	56	33.3	1019	1	POL_HV2BE
32	56	33.3	1022	1	POL_HV2BE
33	56	33.3	1054	1	POL_HV2BE

34	56	33.3	1056	1	POL_SIVM1	P05896 simian immu
35	55	32.7	1124	1	POL_FIVPE	P16088 feline immu
36	55	32.7	1124	1	POL_FIVSD	P19028 feline immu
37	53	31.5	1009	1	POL_SIVGB	P22382 simian immu
38	53	31.5	1034	1	POL_HV2CA	P24107 human immu
39	53	31.5	1035	1	POL_HV2CA	P05962 human immu
40	53	31.5	1036	1	POL_HV2CA	P04584 human immu
41	53	31.5	1036	1	POL_HV2CA	P15833 human immu
42	53	31.5	1165	1	POL_GALV	P21414 gibbon ape
43	52	31.0	481	1	GSHI_CIOAB	Q971V1 clostridium
44	51	30.4	206	1	URK_LACIA	Q9CF21 lactococcus
45	51	30.4	1196	1	POL_MIVRD	P11227 radiation m

ALIGNMENTS

RESULT 1
ID POL_HV1A2 STANDARD; PRT; 1003 AA.
AC P03369;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
RA Stempien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
RA Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
RT (ARV-2).";
RL Science 227:484-492(1985).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 57-155.
RX MEDLINE=89346747; PubMed=2548279;
RA Wlodawer A., Miller M., Jaskolski M., Sathyanarayana B.K.,
RA Baldwin E., Weber I.T., Selk L.M., Clawson L., Schneider J.,
RA Kent S.B.H.;
RT "Conserved folding in retroviral proteases: crystal structure of a
RT synthetic HIV-1 protease";
RL Science 245:616-621(1989).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS) OF 57-155 OF COMPLEX WITH INH.
RA Abbenante G., March D.R., Bergman D.A., Hunt P.A.,
RA Garham B., Dancer R.J., Martin J.L., Fairlie D.P.,
RT "Regioselective structural and functional mimicry of peptides - design
RT of hydrolytically-stable cyclic peptidomimetic inhibitors of HIV-1
RT protease";
RL J. Am. Chem. Soc. 117:10220-10226(1995).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS) OF 57-155 OF COMPLEX WITH INH.
RA March D.R., Abbenante G., Bergman D.A., Brinkworth R.I.,
RA Wickramasinghe W., Begun J., Martin J.L., Fairlie D.P.;
RT "Substrate-based cyclic peptidomimetics of Phe-1le-Val that inhibit
RT HIV-1 protease using a novel enzyme-binding mode";
RL J. Am. Chem. Soc. 118:3375-3379(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 57-155.
RX MEDLINE=96438794; PubMed=8841139;
RA Rose R.B., Craik C.S., Douglas N.L., Stroud R.M.;
RT "Three-dimensional structures of HIV-1 and SIV protease product
RT complexes";
RL Biochemistry 35:12933-12944(1996).
CC -I- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
CC and PI' variable, but often Pro.
CC -I- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-

phosphomonoester.
 -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + (DNA)(N).
 -1- PTM: Cleavage sites that yield the mature proteins remain to be determined.
 -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO KNOWN AS THE RETROPEPSIN FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch>).

 DR EMBL: K02007; AAB59876.1; -.
 DR PIR: A03968; GNWMA2.
 DR PDB: 3HPV; 15-JAN-90.
 DR PDB: 1CPI; 08-MAR-96.
 DR PDB: 1MTR; 01-AUG-96.
 DR PDB: 1YTG; 12-MAR-97.
 DR PDB: 1YTH; 12-MAR-97.
 DR PDB: 1AID; 15-OCT-97.
 DR PDB: 1B6T; 07-JAN-00.
 DR PDB: 1B6K; 07-JAN-00.
 DR PDB: 1B6L; 07-JAN-00.
 DR PDB: 1B6M; 07-JAN-00.
 DR PDB: 1B6N; 07-JAN-00.
 DR PDB: 1B6P; 07-JAN-00.
 DR PDB: 1D4K; 11-OCT-00.
 DR PDB: 1D4L; 01-NOV-00.
 DR PDB: 1F7A; 27-JUN-01.
 DR PDB: 1KJ4; 06-MAR-02.
 DR PDB: 1KJ7; 06-MAR-02.
 DR PDB: 1KJF; 06-MAR-02.
 DR PDB: 1KJG; 06-MAR-02.
 DR PDB: 1KJH; 06-MAR-02.
 DR PDB: 1KZK; 03-APR-02.
 DR PDB: 1MT7; 07-JAN-03.
 DR PDB: 1MT8; 07-JAN-03.
 DR PDB: 1MT9; 07-JAN-03.
 DR PDB: 1MT3; 07-JAN-03.
 DR PDB: 1N49; 07-JAN-03.
 DR PDB: 2AID; 15-OCT-97.
 DR PDB: 3AID; 17-SEP-97.
 DR PDB: 4HVP; 15-OCT-92.
 DR PDB: 8HVP; 31-OCT-93.
 DR HIV: K02007; POLSF2.
 DR MEROPS: A02.001; -.
 DR InterPro: IPR001969; Asparticase_AS.
 DR InterPro: IPR001037; Integrase_C.
 DR InterPro: IPR003308; Integrase_Zn.
 DR InterPro: IPR009007; Pept_A acid.
 DR InterPro: IPR001995; Peptidase_A2.
 DR InterPro: IPR002156; RNaseH.
 DR InterPro: IPR001584; Rve.
 DR InterPro: IPR00477; RVTse.
 DR Pfam: PF00552; Integrase_1.
 DR Pfam: PF02022; Integrase_Zn_1.
 DR Pfam: PF00075; RNaseH_1.
 DR Pfam: PF00665; Rve_1.
 DR Pfam: PF00077; VTP_1.
 DR Pfam: PF00078; VTP_1.
 DR PROSITE: PS00141; ASP_PROTASE_1.
 DR PROSITE: PS00175; ASP_PROTASE_1.
 DR AIDS: Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease; Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
 KM CHAIN 57 155
 FT ACT SITE 81 155
 FT STRAND 58 59
 FT STRAND 66 71
 FT STRAND 66 71
 BY SIMILARITY.

FT TURN 72 73
 FT STRAND 74 80
 FT TURN 82 83
 FT STRAND 88 89
 FT STRAND 99 105
 FT TURN 106 107
 FT STRAND 108 122
 FT TURN 123 124
 FT STRAND 125 133
 FT STRAND 139 141
 FT HELIX 143 146
 FT TURN 147 150
 FT STRAND 152 154
 SO SEQUENCE 1003 AA; 113723 MW; 1519A67242219CB3 CRC64;
 Query Match 44.6%; Score 75; DB 1; Length 1003;
 Best Local Similarity 87.5%; Pred. No. 0.0037;
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 17 RGNPDIVYQYMDL 32
 Db 327 RKNPDIVYQYMDL 342
 RESULT 2
 ID POL_HVH2 STANDARD; PRT; 1003 AA.
 AC P04585; O09777; Q9WUC5;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE POL. polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16); DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN POL.
 OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 NC NCBI_Taxid=11706;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6729196; PubMed=3040055;
 RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S., Gallo R.C., Wong-Staal F.;
 RT "Complete nucleotide sequences of functional clones of the AIDS virus";
 RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
 RN [2]
 RP REVISIONS.
 RA Ogata N., Alter H.J., Miller R.H., Purcell R.H.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chappay C.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 57-155.
 RX MEDLINE=99043699; PubMed=9827997;
 RA Kervinen J., Lubkowski J., Zdanov A., Bharti D., Dunn B.M., Hui K.Y., Powell D.J., Kay J., Mlodner A., Gustchina A.;
 RT "Toward a universal inhibitor of retroviral proteases: comparative analysis of the interactions of LP-130 complexed with proteases from HIV-1, FIV, and EIAV";
 RL Protein Sci. 7:2314-2323(1998).
 RN [5]
 RP STRUCTURE BY NMR OF 57-155.
 RX MEDLINE=9702126; PubMed=868486;
 RA Yamazaki T., Hinck A.P., Wang Y.-X., Nicholson L.K., Torchia D.A., Lam P.Y.S.;
 RT "Three-dimensional solution structure of the HIV-1 protease complexed with DMP323, a novel cyclic urea-type inhibitor, determined by nuclear magnetic resonance spectroscopy";
 RL Protein Sci. 5:495-506(1996).
 RN [6]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 156-595.
RX MEDLINE=96097398; PubMed=8535785;
RA Ren J., Enouf R.M., Hopkins A.L., Ross C., Jones E.Y., Stammers D.K.,
RA Stuart D.I.;
RT "The structure of HIV-1 reverse transcriptase complexed with
RT 9-chloro-TIBO: lessons for inhibitor design.";
RL Structure 3:915-926(1995).
[7]
RN X-RAY CRYSTALLOGRAPHY (2.55 ANGSTROMS) OF 156-595.
RX MEDLINE=96208551; PubMed=8648598;
RA Hopkins A.L., Ren J., Enouf R.M., Willcox B.E., Jones E.Y., Ross C.,
RA Miyasaka T., Walker R.T., Tanaka H., Stammers D.K., Stuart D.I.;
RT "Complexes of HIV-1 reverse transcriptase with inhibitors of the HEPT
RT series reveal conformational changes relevant to the design of potent
RT non-nucleoside inhibitors.";
RL J. Med. Chem. 39:1589-1600(1996).
[8]
RN X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 157-694.
RX MEDLINE=97268683; PubMed=9108091;
RA Enouf R.M., Ren J., Hopkins A.L., Ross C.K., Jones E.Y.,
RA Stammers D.K., Stuart D.I.;
RT "Unique features in the structure of the complex between HIV-1
RT reverse transcriptase and the bis(heteroaryl)piperazine (BHAP)
RT U-90152 explain resistance mutations for this nonnucleoside
RT inhibitor";
RL Proc. Natl. Acad. Sci. U.S.A. 94:3984-3989(1997).
[9]
RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 157-694.
RX MEDLINE=98356189; PubMed=9689112;
RA Ren J., Enouf R.M., Hopkins A.L., Jones E.Y., Kirby I., Keeling J.,
RA Ross C.K., Larder B.A., Stuart D.I., Stammers D.K.;
RT "3'-Azido-3'-deoxythymidine drug resistance mutations in HIV-1
RT reverse transcriptase can induce long range conformational changes.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:9518-9523(1998).
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
CC and PI' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; K03455; AAB50259.1; ALT_INIT.
DR EMBL; AF033819; AAC82598.2; -.
DR PDB; 1REV; 14-OCT-96.
DR PDB; 1RT1; 21-APR-97.
DR PDB; 1RT2; 21-APR-97.
DR PDB; 1RT3; 16-FEB-99.
DR PDB; 1RTH; 03-APR-96.
DR PDB; 1RTJ; 03-APR-96.
DR PDB; 1RTJ; 03-APR-96.
DR PDB; 1BEV; 17-AUG-96.
DR PDB; 1BVG; 17-AUG-96.
DR PDB; 1KIM; 18-MAR-98.
DR PDB; 1A30; 29-APR-98.
DR PDB; 1ODY; 16-FEB-99.
DR PDB; 1BV7; 14-JAN-00.
DR PDB; 1BV9; 12-JAN-00.
DR PDB; 1BWA; 12-JAN-00.
DR PDB; 1BMB; 12-JAN-00.
DR PDB; 1COT; 19-JUL-00.
DR PDB; 1COU; 19-JUL-00.

DR PDB; 1CUB; 21-JUL-00.
DR PDB; 1CIC; 06-SEP-00.
DR PDB; 1DMP; 12-NOV-97.
DR PDB; 1DTQ; 20-MAR-00.
DR PDB; 1DTT; 02-APR-00.
DR PDB; 1EP4; 27-SEP-00.
DR PDB; 1EX4; 26-JUL-00.
DR PDB; 1EXQ; 03-NOV-00.
DR PDB; 1FK9; 28-FEB-01.
DR PDB; 1FKO; 28-FEB-01.
DR PDB; 1FKP; 03-NOV-00.
DR PDB; 1HVA; 06-APR-99.
DR PDB; 1HYR; 15-MAY-95.
DR PDB; 1HWR; 23-MAR-99.
DR PDB; 1HXB; 12-MAR-97.
DR PDB; 1JXH; 03-OCT-01.
DR PDB; 1JUA; 03-OCT-01.
DR PDB; 1JUB; 03-OCT-01.
DR PDB; 1JUC; 03-OCT-01.
DR PDB; 1JUF; 03-OCT-01.
DR PDB; 1JUG; 03-OCT-01.
DR PDB; 1JUG; 22-AUG-01.
DR PDB; 1JW0; 30-OCT-02.
DR PDB; 1JW2; 30-OCT-02.
DR PDB; 1JWC; 30-OCT-02.
DR PDB; 1JWE; 30-OCT-02.
DR PDB; 1JWF; 30-OCT-02.
DR PDB; 1JER; 15-APR-98.
DR PDB; 1JES; 15-APR-98.
DR PDB; 1MET; 15-APR-98.
DR PDB; 1MEU; 15-APR-98.
DR PDB; 1OIW; 18-FEB-03.
DR PDB; 1ODW; 01-APR-97.
DR PDB; 1OBR; 15-OCT-97.
DR PDB; 1OBS; 15-OCT-97.
DR PDB; 1OBT; 15-OCT-97.
DR PDB; 1OBU; 15-OCT-97.
DR PDB; 1RT4; 29-JUL-99.
DR PDB; 1RT5; 29-JUL-99.
DR PDB; 1RT6; 29-JUL-99.
DR PDB; 1RT7; 29-JUL-99.
DR PDB; 1RTD; 12-JAN-00.
DR PDB; 1VRT; 03-APR-96.
DR PDB; 3PHV; 15-JAN-92.
DR HIV; K03455; POLSHXB2.
DR MEROPS; A02.001.1.1;
DR InterPro; IPR001969; Asparticase_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR00308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR00477; RVTse.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; rnaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvc; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS01075; ASP_PROT_RETROV; 1.
DR AIDS; Polyprotein; Hydroxylase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 57 155
FT ACT_SITE 81 81
FT STRAND 58 59
FT STRAND 66 71
FT TURN 72 73
FT STRAND 74 80
FT TURN 82 83
FT STRAND 88 89
FT STRAND 99 104

```

FT STRAND 109 122
FT STRAND 126 133
FT STRAND 140 141
FT HELIX 143 146
FT TURN 147 150
FT STRAND 152 154
FT STRAND 167 167

Query Match 44.6%; Score 75; DB 1; Length 1003;
Best Local Similarity 87.5%; Pred. No. 0.0037;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGNPDIYQYMDL 32
Db 327 RKONPDIYQYMDL 342

RESULT 3
POL_HV1NS STANDARD; PRT; 1003 AA.
ID POL_HV1NS
AC P12497;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polypeptide [contains: Protease (retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
POL.
GN Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).
OS Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OC NCBI_Taxid=11698;
RX MEDLINE=90354401; PubMed=2201682;
RA Fitzergerald P.M.D., McKeever B.M., van Middlesworth J.F.,
RA Springer J.P., Heimbach J.C., Leu C.-T., Herber W.K., Dixon R.A.F.,
RA Darke P.L.;
RT "Crystallographic analysis of a complex between human
immunodeficiency virus type 1 protease and acetyl-peptstatin at 2.0-A
resolution.";
RL J. Biol. Chem. 265:14209-14219(1990).
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
and PI' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomononucleoside.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA} (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M19921; AAA44988.1; -.
CC PDB; 5HVP; 15-OCT-91.
CC PDB; 4PHV; 31-OCT-93.
CC PDB; 1B9F; 19-JUL-99.
CC PDB; 1BHU; 28-OCT-98.
CC PDB; 1B14; 18-NOV-98.
CC PDB; 1B1S; 16-SEP-98.
CC PDB; 1B1U; 16-SEP-98.
CC PDB; 1WUB; 13-MAY-98.
CC PDB; 1WUD; 13-MAY-98.

```

```

DR PDB; 21TG; 12-MAR-97.
DR PDB; 9HVP; 15-JUL-92.
DR HIV; M19921; POLSML43.
DR MEROPS; A02.001; -.
DR InterPro; IPR001699; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Peptidase_Zn.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF02022; Integrase_Zn_1.
DR Pfam; PF00075; RNaseH_1.
DR Pfam; PF00665; Rve_1.
DR Pfam; PF00077; RVD_1.
DR Pfam; PF00078; RVT_1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR AIDS; Polypeptide, Hydrolyase, Aspartyl protease; Endonuclease;
Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
KW CHAIN 57 155
FT ACT SITE 81 81
FT STRAND 58 60
FT STRAND 66 71
FT TURN 72 73
FT STRAND 74 80
FT TURN 82 83
FT STRAND 87 90
FT STRAND 99 105
FT TURN 106 107
FT STRAND 108 122
FT TURN 123 124
FT STRAND 125 134
FT STRAND 140 141
FT HELIX 143 146
FT HELIX 147 149
FT TURN 150 150
FT STRAND 152 154
SQ SEQUENCE 1003 AA; 113535 MW; EB9CDBA30E114A55 CRC64;

Query Match 44.6%; Score 75; DB 1; Length 1003;
Best Local Similarity 87.5%; Pred. No. 0.0037;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGNPDIYQYMDL 32
Db 327 RKONPDIYQYMDL 342

RESULT 4
POL_HV10Y STANDARD; PRT; 1003 AA.
ID POL_HV10Y
AC P20892;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL polypeptide [contains: Protease (retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (OVI isolate) (HIV-1).
OS Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OC NCBI_Taxid=11699;
RX MEDLINE=90148544; PubMed=2559749;
RA Hueb T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Main-Hobson S.;
RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
individual presenting an atypical western blot.";
RL AIDS 3:707-715(1989).
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
and PI' variable, but often Pro.

```



```
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A
CC HEALTHY GABONESE INDIVIDUAL.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
CC PATIENT IN 1984.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M26727; AAA83392.1; -.
CC HSSP: P03366; 1RVL.
CC HIV: M26727; POL$OYL.
CC MEROPS: A02.001; -.
CC InterPro: IPR001969; Aspprotease_AS.
CC InterPro: IPR001037; Integrase_C.
CC InterPro: IPR003308; Integrase_Zn.
CC InterPro: IPR009007; Pept_A_acid.
CC InterPro: IPR001995; Peptidase_A2.
CC InterPro: IPR002156; RNaseH.
CC InterPro: IPR001584; Rve.
CC InterPro: IPR000477; RVTse.
CC InterPro: IPR00552; Integrase_1.
CC Pfam: PF02022; Integrase_Zn; 1.
CC Pfam: PF00075; RNaseH; 1.
CC Pfam: PF00665; rve; 1.
CC Pfam: PF00077; rvp; 1.
CC Pfam: PF00078; rvt; 1.
CC PROSITE: PS00141; ASP_PROTASE. 1.
CC PROSITE: PS50175; ASP_PROT_RETROV. 1.
CC AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
CC Nuclease; Transferase; RNA-directed DNA polymerase.
CC FT CHAIN 57 155 BY SIMILARITY.
CC ACT SITE 81 81 IN-FRAME TERMINATION CODON.
CC SITE 565 565 IN-FRAME TERMINATION CODON.
CC SEQUENCE 1003 AA; 113718 MW; AFB997A0ED88A98 CRC64;
SQ
Query Match 44.6%; Score 75; DB 1; Length 1003;
Best Local Similarity 87.5%; Pred. No. 0.0037;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 17 RGRNPDIYQYMDL 32
Db 327 RKQNPDIYQYMDL 342
RESULT 5
POL_HVIMN STANDARD; PRT; 1006 AA.
AC P05961;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL polyprotein [Contains: Protease (Retriopepsin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)] .
GN POL.
OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCB1_TaxID=11696;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=88219542; PubMed=3369091; Aldovini A., Collalti E.,
RA Gurgo C., Guo H.-G., Franchini G., Gallo R.C., Reltz M. Jr.,
RA Farrell K., Wong-Staal F., Gallo R.C., Reltz M. Jr.,
RA "Envelope sequences of two new United States HIV-1 isolates.";
```

```
RL Virology 164:531-536(1988).
CC -1- CATALYTIC ACTIVITY: Specific for a PL residue that is hydrophobic,
CC and PL' variable, but often pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
CC PATIENT IN 1984.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M17449; -, NOT ANNOTATED_CDS.
CC HSSP: P03366; 1RVL.
CC HIV: M17449; POL$NMN.
CC MEROPS: A02.001; -.
CC InterPro: IPR001969; Aspprotease_AS.
CC InterPro: IPR001037; Integrase_C.
CC InterPro: IPR003308; Integrase_Zn.
CC InterPro: IPR009007; Pept_A_acid.
CC InterPro: IPR001995; Peptidase_A2.
CC InterPro: IPR002156; RNaseH.
CC InterPro: IPR001584; Rve.
CC InterPro: IPR000477; RVTse.
CC InterPro: IPR00552; Integrase_1.
CC Pfam: PF02022; Integrase_Zn; 1.
CC Pfam: PF00075; RNaseH; 1.
CC Pfam: PF00665; rve; 1.
CC Pfam: PF00077; rvp; 1.
CC Pfam: PF00078; rvt; 1.
CC PROSITE: PS00141; ASP_PROTASE. 1.
CC PROSITE: PS50175; ASP_PROT_RETROV. 1.
CC AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
CC Nuclease; Transferase; RNA-directed DNA polymerase.
CC FT CHAIN 60 158 BY SIMILARITY.
CC ACT SITE 84 84 IN-FRAME TERMINATION CODON.
CC SITE 565 565 IN-FRAME TERMINATION CODON.
CC SEQUENCE 1006 AA; 113860 MW; 460C50DF92AF9B3 CRC64;
SQ
Query Match 44.6%; Score 75; DB 1; Length 1006;
Best Local Similarity 87.5%; Pred. No. 0.0037;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 17 RGRNPDIYQYMDL 32
Db 330 RKQNPDIYQYMDL 345
RESULT 6
POL_HVIB5 STANDARD; PRT; 1015 AA.
AC P04587;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [Contains: Protease (Retriopepsin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)] .
GN POL.
OS Human immunodeficiency virus type 1 (BHS isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCB1_TaxID=11682;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RX MEDLINE=55111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rataleki J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J., Vamvak T.S., Grunyard J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RA "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284(1985).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
CC and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoster.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- pmv: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02012; AAA44653.1; -.
DR PDB; 1BDL; 12-AUG-98.
DR PDB; 1BDQ; 12-AUG-98.
DR PDB; 1BDR; 14-OCT-98.
DR PDB; 1FEU; 01-JUN-01.
DR PDB; 1FE0; 01-JUN-01.
DR PDB; 1FEF; 01-JUN-01.
DR PDB; 1FFI; 01-JUN-01.
DR PDB; 1EG6; 01-JUN-01.
DR PDB; 1EG8; 01-JUN-01.
DR PDB; 1EGC; 01-JUN-01.
DR PDB; 1GK2; 16-NOV-01.
DR PDB; 1HPV; 31-MAR-95.
DR PDB; 1HTE; 31-JUL-94.
DR PDB; 1HVJ; 30-APR-94.
DR PDB; 1HVP; 15-JAN-95.
DR PDB; 1HVS; 14-FEB-95.
DR PDB; 1KI7; 10-JUL-02.
DR PDB; 1KIU; 10-JUL-02.
DR PDB; 1K2B; 10-JUL-02.
DR PDB; 1K2C; 10-JUL-02.
DR PDB; 1ODX; 01-APR-97.
DR PDB; 1TCX; 07-DEC-96.
DR PDB; 1W4E; 16-DEC-98.
DR PDB; 1W6F; 16-DEC-98.
DR PDB; 2BPV; 23-FEB-99.
DR PDB; 2BPW; 23-FEB-99.
DR PDB; 2BPX; 23-FEB-99.
DR PDB; 2BPY; 23-FEB-99.
DR PDB; 2BPZ; 23-FEB-99.
DR HIV; K02012; POLS9H5.
DR MEROPS; A02.001; -.
DR InterPro; IPR001367; Aspprotease_A5.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00552; integrase; 1.
DR Pfam; PF02022; integrase; 1.
DR Pfam; PF00075; rnaseh; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; ryl; 1.

```

```

DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PSS0175; ASP_PROT_RETROV; 1.
KW AIDS; Polyprotein; Hydrolysis; Aspartyl protease; Endonuclease.
RN Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 69 167
FT ACT_SITE 93 193
FT SEQUENCE 1015 AA; 115015 MW; F79FOCB4A1A92CEE CMC64;
SQ
Query Match 44.6%; Score 75; DB 1; Length 1015;
Best Local Similarity 87.5%; Pred. No. 0.0038;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 17 RGRNPDIIVTYQYMDL 32
Db 339 RKNPDIIVTYQYMDL 354
POL_HY1BR STANDARD; PRT; 1015 AA.
ID POL_HY1BR
AC POL367;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirinae.
OX NCBI_TaxID=11686;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85099333; PubMed=2981635;
RA Wann-Hobson S., Sonigo P., Dano O., Cole S., Allizon M.;
RL "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).
RN [2]
RP REVISIONS TO 23-35.
RX MEDLINE=86245056; PubMed=2424612;
RA Allizon M., Wann-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus; nucleotide sequence analysis
RT of two isolates from African patients.";
RL Cell 46:63-74(1986).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX MEDLINE=92190341; PubMed=1796632;
RA Spinnelli S., Liu Q.Z., Alzari P.M., Harel P.H., Poljak R.J.;
RT "The three-dimensional structure of the aspartyl protease from the
RT HIV-1 isolate BRU."
RL Biochimie 73:1391-1396(1991).
CC -I- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
CC and P1' variable, but often Pro.
CC -I- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -I- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [dNA] (N).
CC -I- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; KO2013; -; NOT ANNOTATED_CDS.
DR PDB; 1HHB; 15-OCT-92.
DR PDB; 1A8G; 15-JUL-98.
DR PDB; 1A94; 16-FEB-99.

```

```
DR PDB: 1D4S; 12-OCT-99.
DR PDB: 1D4Y; 13-OCT-99.
DR PDB: 1DAZ; 03-MAY-00.
DR PDB: 1DIF; 08-MAR-96.
DR PDB: 1HOS; 31-OCT-93.
DR PDB: 1HFO; 21-APR-97.
DR PDB: 1HPS; 15-JAN-95.
DR PDB: 1HPX; 08-MAR-96.
DR PDB: 1HSG; 03-APR-96.
DR PDB: 1HTE; 31-JUL-94.
DR PDB: 1HVL; 30-APR-94.
DR PDB: 1UPJ; 14-OCT-96.
DR PDB: 2UPJ; 14-OCT-96.
DR PDB: 7UPJ; 21-APR-97.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asparticase_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvt; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROT_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR AIDS; Polypeptide; Hydroxylase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 69 167
FT ACT SITE 93 93 BY SIMILARITY.
FT STRAND 78 82
FT TURN 87 92
FT TURN 94 95
FT STRAND 100 102
FT TURN 111 117
FT TURN 118 119
FT STRAND 120 134
FT TURN 135 136
FT STRAND 137 146
FT STRAND 152 153
FT HELIX 155 161
FT TURN 162 162
SQ SEQUENCE 1015 AA; 115031 MW; 164702F074A84394 CRC64;

Query Match 44.6%; Score 75; DB 1; Length 1015;
Best Local Similarity 87.5%; Pred. No. 0.0038;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYIYQYMDL 32
DB 339 RKQNPDIYIYQYMDL 354
```

```
RA MEDLINE=85111157; PubMed=2982104;
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,
RA Capon D.J.;
RT "Nucleic acid structure and expression of the human
RL Aids/lymphadenopathy retrovirus.";
RL Nature 313:450-458 (1985).
RN [2]
RP REVISION.
RA Muesing M.A.;
RL Submitted (xxx-1987) to the HIV data bank.
CC -! CATALYTIC ACTIVITY: Specific for a PL residue that is hydrophobic,
CC and PL variable, but often Pro.
CC -! CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -! CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {dNp}(N).
CC -! PM: cleavage sites that yield the mature proteins remain to be
CC determined.
CC -! SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; K02083; AB59867.1; -.
DR EMBL; X01762; -; NOT ANNOTATED_CDS.
DR PIR; A03967; GNVWVL.
DR PDB; 1A9M; 17-JUN-98.
DR PDB; 1AJV; 20-AUG-97.
DR PDB; 1AJX; 17-SEP-97.
DR PDB; 1G35; 06-JUN-01.
DR PDB; 1GNM; 08-NOV-96.
DR PDB; 1HTG; 31-JUL-94.
DR PDB; 1HVI; 30-APR-94.
DR PDB; 1KIH; 06-MAR-02.
DR PDB; 1NPV; 04-FEB-03.
DR PDB; 1NPW; 04-FEB-03.
DR HIV; K02083; POLSPV22.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asparticase_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvt; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR AIDS; Polypeptide; Hydroxylase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 69 167
FT ACT SITE 93 93 BY SIMILARITY.
SQ SEQUENCE 1015 AA; 115090 MW; 51529D18EAC2AF89 CRC64;

Query Match 44.6%; Score 75; DB 1; Length 1015;
Best Local Similarity 87.5%; Pred. No. 0.0038;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYIYQYMDL 32
DB 339 RKQNPDIYIYQYMDL 354
```

```

RESULT 9
POL_HV1JR STANDARD; PRT: 1007 AA.
ID POL_HV1JR
AC P35963;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL polyprotein [contains: Protease (Retrovirsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=11688;
RN [1]
RP SEQUENCE FROM N.A.
RA Koyanagi S., Chen I.S.Y.;
RL Submitted (DEC-1988) to the HIV data bank.
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
CC and PI' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA](N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M38429; AAB03745.1; -.
CC HSSP; P03366; 1HMV.
CC DR HIV; M38429; POL$JRCSE.
CC DR MEROPS; A02.001; -.
CC DR InterPro; IPR001969; Asprotease_AS.
CC DR InterPro; IPR001037; Integrase_C.
CC DR InterPro; IPR003308; Integrase_Zn.
CC DR InterPro; IPR009007; Pept_A_acid.
CC DR InterPro; IPR001995; Peptidase_A2.
CC DR InterPro; IPR002156; RNaseH.
CC DR InterPro; IPR001584; Rve.
CC DR InterPro; IPR000477; RYase.
CC DR Pfam; PF00552; Integrase; 1.
CC DR Pfam; PF02022; Integrase_Zn; 1.
CC DR Pfam; PF00075; naseH; 1.
CC DR Pfam; PF00665; rve; 1.
CC DR Pfam; PF00077; rvp; 1.
CC DR Pfam; PF00078; rvc; 1.
CC DR PROSITE; PS00141; ASP_PROTEASE; 1.
CC DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
CC DR AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
CC Nuclease; Transferase; RNA-directed DNA polymerase.
CC FT CHAIN 61 159
CC ACT SITE 85 85 BY SIMILARITY.
CC SEQUENCE 1007 AA; 114081 MW; B123656C2EP3411A CRC64;
Query Match 44.0%; Score 74; DB 1; Length 1007;
Best Local Similarity 81.2%; Pred. No. 0.0053;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
DB 331 RKNPDIIVYQWDDL 32
17 RKNPDIIVYQWDDL 32
331 RKNPDIIVYQWDDL 346

```

```

RESULT 10
POL_HV1Y2 STANDARD; PRT: 1003 AA.
ID POL_HV1Y2
AC P35963;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [contains: Protease (Retrovirsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC Viruses; Retrovirdae; Lentivirus.
OX NCBI_TaxID=36377;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93021387; PubMed=1404605;
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
RA Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological
RT properties of human immunodeficiency virus type 1 in vivo: evidence
RT for limited defectiveness and complementation."
RL J. Virol. 66:6587-6600(1992).
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
CC and PI' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA](N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M93258; -; NOT_ANNOTATED_CDS.
CC DR PIR; B44001; B44001.
CC DR PDB; 1K6C; 06-FEB-02.
CC DR PDB; 1K6P; 06-FEB-02.
CC DR PDB; 1K6T; 06-FEB-02.
CC DR PDB; 1K6V; 06-FEB-02.
CC DR PDB; 1WJA; 13-MAY-98.
CC DR PDB; 1WJC; 13-MAY-98.
CC DR MEROPS; A02.001; -.
CC DR InterPro; IPR001969; Asprotease_AS.
CC DR InterPro; IPR001037; Integrase_C.
CC DR InterPro; IPR003308; Integrase_Zn.
CC DR InterPro; IPR009007; Pept_A_acid.
CC DR InterPro; IPR001995; Peptidase_A2.
CC DR InterPro; IPR002156; RNaseH.
CC DR InterPro; IPR001584; Rve.
CC DR InterPro; IPR000477; RYase.
CC DR Pfam; PF00552; Integrase; 1.
CC DR Pfam; PF02022; Integrase_Zn; 1.
CC DR Pfam; PF00075; naseH; 1.
CC DR Pfam; PF00665; rve; 1.
CC DR Pfam; PF00077; rvp; 1.
CC DR Pfam; PF00078; rvc; 1.
CC DR PROSITE; PS00141; ASP_PROTEASE; 1.
CC DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
CC DR AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
CC Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
CC FT CHAIN 57 155
CC ACT SITE 81 81 BY SIMILARITY.
CC SEQUENCE 1003 AA; 113794 MW; D2F7187FE4EBF49 CRC64;
Query Match 43.5%; Score 73; DB 1; Length 1003;

```

Best Local Similarity 81.2%; Pred. No. 0.0075;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYQYMDL 32
| : : : : :
Db 327 RKQNPDIYQYMDL 342

```

RESULT 11
POL_HV1MA STANDARD; PRT; 1002 AA.
AC P04568; Q79582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (MAL isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11697;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients.";
RL Cell 46:63-74(1986).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA} (N).
CC -1- P1N: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04415; CAA28012.1; -.
DR EMBL; A07116; CAA00619.1; -.
DR PDB; 1HHJ; 31-OCT-93.
DR HIV; K03456; POLSMAL.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RYase.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; Rve; 1.
DR Pfam; PF00077; Tvp; 1.
DR Pfam; PF00078; Tvp; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR AIDS; Polypoteins; Hydrolyase; Aspartyl protease; Endonuclease;
RNase; Transferrase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 56 154
ACT_SITE 80 154 BY SIMILARITY.
SEQUENCE 1002 AA; 113537 MW; 569A29D69AC6AC5 CRC64;

```

Query Match 42.9%; Score 72; DB 1; Length 1002;
Best Local Similarity 81.2%; Pred. No. 0.011;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYQYMDL 32
| : : : : :
Db 326 RTKNPDIYQYMDL 341

```

RESULT 12
POL_HV1U4 STANDARD; PRT; 1002 AA.
AC P24740;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (strain Ugandan / isolate U455)
(HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11703;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91090981; PubMed=2265025;
RA Oram J.D., Downing R.G., Roff M., Clegg J.C.S., Serwadda D.,
Carwell J.W.;
RT "Nucleotide sequence of a Ugandan HIV-1 provirus reveals genetic
diversity from other HIV-1 isolates.";
RL AIDS Res. Hum. Retroviruses 6:1073-1078(1990).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA} (N).
CC -1- P1N: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62320; AAA75019.1; -.
DR PDB; 1E27; 12-SEP-00.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RYase.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; Rve; 1.
DR Pfam; PF00077; Tvp; 1.
DR Pfam; PF00078; Tvp; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR AIDS; Polypoteins; Hydrolyase; Aspartyl protease; Endonuclease;
RNase; Transferrase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 56 154
PROTEASE.

```


Db 351 RKNDPDIYQYMDL 366

```

RESULT 15
POL_HVIND
ID POL_HVIND STANDARD; PRT; 1002 AA.
AC P18602;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL polyprotein [contains: Protease (retropepsin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11695;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90034200; PubMed=2806917;
RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,
RA Hampe A., Chermann J.C.;
RT "Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the
RT human immunodeficiency virus.";
RL Gene 81:275-284(1989).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
CC and P1 variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- P.TM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M27323; AAA44869.1; -.
CC DR PIR; JQ0067; GNLJND.
CC DR HSSP; P03366; IHMV.
CC DR HIV; M27323; POL$NDK.
CC DR MEROPS; A02.001; -.
CC DR InterPro; IPR001969; Asprotease_AS.
CC DR InterPro; IPR001037; Integrase_C.
CC DR InterPro; IPR003308; Integrase_Zn.
CC DR InterPro; IPR009007; Pept_A_acid.
CC DR InterPro; IPR001995; Peptidase_A2.
CC DR InterPro; IPR002155; RNaseH.
CC DR InterPro; IPR001584; Rve.
CC DR InterPro; IPR000477; RVTse.
CC DR Pfam; PF00552; Integrase_1.
CC DR Pfam; PF02022; Integrase_Zn; 1.
CC DR Pfam; PF00075; rnaseH; 1.
CC DR Pfam; PF00665; rve; 1.
CC DR Pfam; PF00077; rvp; 1.
CC DR Pfam; PF00078; rvc; 1.
CC DR PROSITE; PS00141; ASP_PROTEASE; 1.
CC DR PROSITE; PSS0175; ASP_PROT_RETROV; 1.
CC KM AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
CC Nuclease; Transferase; RNA-directed DNA polymerase.
CC CHAIN 56 154
FT ACT_SITE 80 80 BY SIMILARITY.
SQ SEQUENCE 1002 AA; 113621 MW; 5ED866E8C9896CC1 CRC64;

```

Query Match 42.3%; Score 71; DB 1; Length 1002;
 Best Local Similarity 81.2%; Pred. No. 0.015;
 Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 17 RGRNPDIVYQYMDL 32
 Db 326 RKNDPDIYQYMDL 341

Search completed: July 20, 2004, 06:31:01
 Job time : 6.29508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:43 ; Search time 32.2623 Seconds
(without alignments)
312.953 Million cell updates/sec

Title: US-09-673-166b-275
Perfect score: 168
Sequence: 1 GRQYRKANSKPTGITERGRNPDIYQYMDL 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 31518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriophage: *
17: sp_archaeal: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	50.0	328	15	Q9EFP6 human immun
2	84	50.0	328	15	Q9EFC0 human immun
3	84	50.0	328	15	Q9EFP4 human immun
4	84	50.0	328	15	Q9EFC3 human immun
5	84	50.0	328	15	Q9EFC8 human immun
6	84	50.0	328	15	Q9EFP8 human immun
7	84	50.0	328	15	Q9EFP3 human immun
8	84	50.0	328	15	Q9EFP7 human immun
9	84	50.0	328	15	Q9EFP4 human immun
10	84	50.0	328	15	Q9EFP4 human immun
11	84	50.0	328	15	Q9EFP4 human immun
12	84	50.0	328	15	Q9EFP4 human immun
13	84	50.0	328	15	Q9EFP4 human immun
14	84	50.0	328	15	Q9EFP4 human immun
15	84	50.0	328	15	Q9EFP4 human immun
16	84	50.0	341	15	Q9G1L1 human immun

17	82	48.8	220	15	Q37429 human immun
18	82	48.8	220	15	Q37404 human immun
19	82	48.8	308	15	Q805U7 human immun
20	81	48.2	190	15	Q7ZSR6 human immun
21	81	48.2	207	15	Q7SNCO human immun
22	81	48.2	207	15	Q7SNB9 human immun
23	81	48.2	409	15	Q9DPE5 human immun
24	81	48.2	411	15	Q8ALQ7 human immun
25	81	48.2	999	15	Q8UTQ9 human immun
26	80	47.6	85	15	Q9WFI6 human immun
27	80	47.6	328	15	Q9EFC7 human immun
28	80	47.6	330	15	Q8EAB9 human immun
29	79	47.0	163	15	Q9WGP8 human immun
30	79	47.0	190	15	Q7SSJ5 human immun
31	79	47.0	191	15	Q902D9 human immun
32	79	47.0	191	15	Q902D1 human immun
33	79	47.0	192	15	Q9WGP5 human immun
34	79	47.0	193	15	Q8Q229 human immun
35	79	47.0	200	15	Q8AMJ5 human immun
36	79	47.0	201	15	Q9IFC3 human immun
37	79	47.0	203	15	Q93140 human immun
38	79	47.0	206	15	Q7SJZ5 human immun
39	79	47.0	206	15	Q7SJZ0 human immun
40	79	47.0	206	15	Q7SJZ9 human immun
41	79	47.0	207	15	Q7SNH7 human immun
42	79	47.0	207	15	Q7SNG1 human immun
43	79	47.0	207	15	Q7SNC7 human immun
44	79	47.0	210	15	Q8JSP4 human immun
45	79	47.0	210	15	Q8APX9 human immun

ALIGNMENTS

RESULT 1
Q9EFP6 PRELIMINARY: PRT; 328 AA.
ID Q9EFP6
AC Q9EFP6; 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pol protein (fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=062C03;
RX MEDLINE=20408577; PubMed=10952598;
RA Bachele L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Aujay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.U., Spalding D.R., Hollis G., Abremki K.,
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
Failing Efavirenz Combination Therapy.";
RT Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL: AY002129; AAF92284.1; -!
DR GO: GO:0006190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0006233; F:peptidase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Asparticase_AS.
DR InterPro: IPR001995; Peptidase_A2.
DR InterPro: IPR009007; Pept_A_acid.
DR Pfam: PF00077; rvt; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.

```
DR PROSITE; PSS0175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT NON_TER 1 1
FT CHAIN 100 >99 PROTEASE.
FT CHAIN 328 >328 REVERSE TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37180 MW; 35CEBB2983A8FC07 CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGRNPDIYIYQYMDL 32
DB 271 RGRNPDIYIYQYMDL 286

RESULT 2
Q9EFQ0 PRELIMINARY; PRT; 328 AA.
AC Q9EFQ0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=062C02;
RX MEDLINE=20408577; Pubmed=10952598;
RA Bacheler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Anjay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Anticreb. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
EMBL; AY002125; AAF92280.1; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Asprotease AS.
DR InterPro; IPR001995; Peptidase A2.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; RVP; 1.
DR PROSITE; PSS0141; ASP_PROTEASE; 1.
DR PROSITE; PSS0175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT NON_TER 1 1
FT CHAIN 100 >99 PROTEASE.
FT CHAIN 328 >328 REVERSE TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37310 MW; 3415A4E02F47FE8 CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGRNPDIYIYQYMDL 32
DB 271 RGRNPDIYIYQYMDL 286
```

```
RESULT 3
Q9EFN4 PRELIMINARY; PRT; 328 AA.
AC Q9EFN4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=091B05;
RX MEDLINE=20408577; Pubmed=10952598;
RA Bacheler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Anjay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Anticreb. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
EMBL; AY002141; AAF92296.1; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Asprotease AS.
DR InterPro; IPR001995; Peptidase A2.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; RVP; 1.
DR PROSITE; PSS0141; ASP_PROTEASE; 1.
DR PROSITE; PSS0175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT NON_TER 1 1
FT CHAIN 100 >99 PROTEASE.
FT CHAIN 328 >328 REVERSE TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37297 MW; A4704362BCB77CD CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGRNPDIYIYQYMDL 32
DB 271 RGRNPDIYIYQYMDL 286

RESULT 4
Q9EFQ9 PRELIMINARY; PRT; 328 AA.
AC Q9EFQ9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
```

RC STRAIN=041F05;
MEDLINE=20408577; PubMed=10952598;
RA Bachelet L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Aufay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL: AY002116; AAF92271.1; -
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0008223; F:peptidase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Aspprotease AS.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept_A_acid.
DR Pfam: PF00078; rvt; 1.
DR Pfam: PF00077; rvt; 1.
DR PROSITE: PS00141; ASP_PROT_RETROV; 1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Hydroxylase; Protease; RNA-directed DNA polymerase;
KW Transferase.
FT CHAIN 1 1
FT CHAIN 100 >328
FT NON_TER 328
SQ SEQUENCE 328 AA; 37181 MW; 776457F54CE3E8A4 CRC64;
Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 17 RGNPDIYQYMDL 32
DB 271 RGNPDIYQYMDL 286
RESULT 5
Q9EF08 PRELIMINARY; PRT; 328 AA.
AC Q9EF08; 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=091B01;
RX MEDLINE=20408577; PubMed=10952598;
RA Bachelet L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Aufay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL: AY002117; AAF92272.1; -
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0008223; F:peptidase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.

DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Aspprotease AS.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept_A_acid.
DR InterPro: IPR000477; rvtse.
DR Pfam: PF00077; rvt; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROT_RETROV; 1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Hydroxylase; Protease; RNA-directed DNA polymerase;
KW Transferase.
FT CHAIN 1 1
FT CHAIN 100 >328
FT NON_TER 328
SQ SEQUENCE 328 AA; 37297 MW; A4704362BCEB77CD CRC64;
Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 17 RGNPDIYQYMDL 32
DB 271 RGNPDIYQYMDL 286

RESULT 6
Q9EFN9 PRELIMINARY; PRT; 328 AA.
AC Q9EFN9; 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=041F03;
RX MEDLINE=20408577; PubMed=10952598;
RA Bachelet L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Aufay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL: AY002136; AAF92291.1; -
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0008223; F:peptidase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Aspprotease AS.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept_A_acid.
DR InterPro: IPR000477; rvtse.
DR Pfam: PF00077; rvt; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROT_RETROV; 1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Hydroxylase; Protease; RNA-directed DNA polymerase;
KW Transferase.
FT CHAIN 1 1
FT CHAIN 100 >328
FT NON_TER 328
SQ SEQUENCE 328 AA; 37288 MW; A4704362BCEB77CD CRC64;
Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 17 RGNPDIYQYMDL 32
DB 271 RGNPDIYQYMDL 286

SEQ SEQUENCE 328 AA; 37218 MW; 31F2676CB94F2EB6 CRC64;
Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 17 RGRNPDIVYQYVMDL 32
DB 271 RGRNPDIVYQYVMDL 286
RESULT 7
ID Q9EFN8 PRELIMINARY; PRT; 328 AA.
AC Q9EFN8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=062C09;
RX MEDLINE=20408577; PubMed=10952598;
RA Bacheler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Auway M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RT Antimicrob. Agents Chemother. 44:2475-2484(2000).
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
EMBL: AY002137; AAF92292.1; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Asparticase AS.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept_A_acid.
DR Pfam: PF00077; rvp; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolyase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT CHAIN 1 1
FT NON_TER 1 >99
FT CHAIN 100 >328
FT NON_TER 328 328
FT SEQUENCE 328 AA; 37252 MW; 36783542BCFF2EAD CRC64;
Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 17 RGRNPDIVYQYVMDL 32
DB 271 RGRNPDIVYQYVMDL 286
RESULT 8
ID Q9EFN3 PRELIMINARY; PRT; 328 AA.
AC Q9EFN3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
GN Pol protein (Fragment).
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=062C06;
RX MEDLINE=20408577; PubMed=10952598;
RA Bacheler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Auway M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RT Antimicrob. Agents Chemother. 44:2475-2484(2000).
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
EMBL: AY002142; AAF92297.1; -.
DR HSSP: P03369; ICPT.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Asparticase AS.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept_A_acid.
DR Pfam: PF00077; rvp; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolyase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT CHAIN 1 1
FT NON_TER 1 >99
FT CHAIN 100 >328
FT NON_TER 328 328
FT SEQUENCE 328 AA; 37281 MW; 1B7ED69EB1BD1B0E CRC64;
Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 17 RGRNPDIVYQYVMDL 32
DB 271 RGRNPDIVYQYVMDL 286
RESULT 9
ID Q9EFN3 PRELIMINARY; PRT; 328 AA.
AC Q9EFN3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=062C07;
RX MEDLINE=20408577; PubMed=10952598;
RA Bacheler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Auway M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients

RT Failing Efavirenz Combination Therapy."
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC EMBL: AY002132; AAF92287.1; -
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Asparticase AS.
DR InterPro: IPR009007; Peptidase A2.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00077; rvp; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROTASE; 1.
DR PROSITE: PS0175; ASP_PROT_RETROV; 1.
KM Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT CHAIN 1 1
FT NON_TER 100 >328 PROTEASE.
FT NON_TER 328 REVERSE TRANSCRIPTASE.
SQ SEQUENCE 328 AA; 37282 MW; 36783542BCB7FE8 CRC64;
Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 17 RGNPDIVITYQYMDL 32
Db 271 RGNPDIVITYQYMDL 286
RESULT 10
Q9EFN7 PRELIMINARY; PRT; 328 AA.
ID Q9EFN7
AC Q9EFN7
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN GN
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=041F01;
RX MEDLINE=20408577; PubMed=10952598;
RA Bacheiler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Antun M., Wang X.V., Ellis D., Becker M.F.,
RA Laust A.L., George H.U., Spalding D.R., Hollis G., Adremski K.,
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy."
RT Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC EMBL: AY002138; AAF92293.1; -
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Asparticase AS.
DR InterPro: IPR009007; Peptidase A2.
DR InterPro: IPR000477; RVTse.

DR Pfam: PF00077; rvp; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0175; ASP_PROT_RETROV; 1.
KM Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT CHAIN 1 1
FT NON_TER 100 >328 PROTEASE.
FT NON_TER 328 REVERSE TRANSCRIPTASE.
SQ SEQUENCE 328 AA; 37282 MW; 36783542BCB7FE8 CRC64;
Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 17 RGNPDIVITYQYMDL 32
Db 271 RGNPDIVITYQYMDL 286
RESULT 11
Q9EP4 PRELIMINARY; PRT; 328 AA.
ID Q9EP4
AC Q9EP4
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN GN
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=070G07;
RX MEDLINE=20408577; PubMed=10952598;
RA Bacheiler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Antun M., Wang X.V., Ellis D., Becker M.F.,
RA Laust A.L., George H.U., Spalding D.R., Hollis G., Adremski K.,
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy."
RT Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC EMBL: AY002131; AAF92286.1; -
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Asparticase AS.
DR InterPro: IPR009007; Peptidase A2.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00077; rvp; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0175; ASP_PROT_RETROV; 1.
KM Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT CHAIN 1 1
FT NON_TER 100 >328 PROTEASE.
FT NON_TER 328 REVERSE TRANSCRIPTASE.
SQ SEQUENCE 328 AA; 37313 MW; B5241227A9BA238D CRC64;
Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY      17 RGNPDIYIYQYMDL 32
DB      271 RGNPDIYIYQYMDL 286

RESULT 12
ID      Q9EF04 PRELIMINARY; PRT; 328 AA.
AC      Q9EF04;
DT      01-MAR-2001 (TREMBLrel. 16, Created)
DT      01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT      01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE      Pol protein (Fragment).
GN      POL.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX      NCB1_Taxid=11676;

RP      SEQUENCE FROM N.A.
RC      STRAIN=091B03;
RX      MEDLINE=20408577; PubMed=10952598;
RA      Bachelet L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA      Krakowski K., Bolling L., Aujay M., Wang X.V., Ellis D., Becker M.F.,
RA      Laust A.L., George H.J., Spalding D.R., Hollis G., Abremski K.;
RT      "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT      Failing Efavirenz Combination Therapy.";
RL      Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC      -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC      KNOWN AS THE RETROPEPSIN FAMILY.

EMBL: AY002121; AAF92276.1; -
DR      GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR      GO: GO:0008233; F:peptidase activity; IEA.
DR      GO: GO:0003723; F:RNA binding; IEA.
DR      GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR      GO: GO:0016740; F:transferase activity; IEA.
DR      GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR      GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR      InterPro: IPR001969; Aspprotease_AS.
DR      InterPro: IPR001995; Peptidase_A2.
DR      InterPro: IPR009007; Pept_A_acid.
DR      InterPro: IPR000477; RVTse.
DR      Pfam: PF00077; rvt; 1.
DR      Pfam: PF00078; rvt; 1.
DR      PROSITE: PS00141; ASP_PROTEASE; 1.
DR      PROSITE: PS50175; ASP_PROT_RETROV; 1.
KW      Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW      Transferase.
FT      CHAIN 1 1
FT      NON_TER 1 >99
FT      CHAIN 100 >328 REVERSE TRANSCRIPTASE.
FT      NON_TER 328 328
SQ      SEQUENCE 328 AA; 37226 MW; 29AE22FD46F8DAA2 CRC64;

Query Match      50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      17 RGNPDIYIYQYMDL 32
DB      271 RGNPDIYIYQYMDL 286

RESULT 13
ID      Q9EF01 PRELIMINARY; PRT; 328 AA.
AC      Q9EF01;
DT      01-MAR-2001 (TREMBLrel. 16, Created)
DT      01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT      01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE      Pol protein (Fragment).
GN      POL.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.

```

```

CX      NCB1_Taxid=11676;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=070G04;
RX      MEDLINE=20408577; PubMed=10952598;
RA      Bachelet L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA      Krakowski K., Bolling L., Aujay M., Wang X.V., Ellis D., Becker M.F.,
RA      Laust A.L., George H.J., Spalding D.R., Hollis G., Abremski K.;
RT      "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT      Failing Efavirenz Combination Therapy.";
RL      Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC      -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC      KNOWN AS THE RETROPEPSIN FAMILY.

EMBL: AY002124; AAF92279.1; -
DR      GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR      GO: GO:0008233; F:peptidase activity; IEA.
DR      GO: GO:0003723; F:RNA binding; IEA.
DR      GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR      GO: GO:0016740; F:transferase activity; IEA.
DR      GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR      GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR      InterPro: IPR001969; Aspprotease_AS.
DR      InterPro: IPR001995; Peptidase_A2.
DR      InterPro: IPR009007; Pept_A_acid.
DR      InterPro: IPR000477; RVTse.
DR      Pfam: PF00077; rvt; 1.
DR      Pfam: PF00078; rvt; 1.
DR      PROSITE: PS00141; ASP_PROTEASE; 1.
DR      PROSITE: PS50175; ASP_PROT_RETROV; 1.
KW      Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW      Transferase.
FT      CHAIN 1 1
FT      NON_TER 1 >99
FT      CHAIN 100 >328 REVERSE TRANSCRIPTASE.
FT      NON_TER 328 328
SQ      SEQUENCE 328 AA; 37297 MW; A4704362BCB877CD CRC64;

Query Match      50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      17 RGNPDIYIYQYMDL 32
DB      271 RGNPDIYIYQYMDL 286

RESULT 14
ID      Q9EF06 PRELIMINARY; PRT; 328 AA.
AC      Q9EF06;
DT      01-MAR-2001 (TREMBLrel. 16, Created)
DT      01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT      01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE      Pol protein (Fragment).
GN      POL.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX      NCB1_Taxid=11676;

RP      SEQUENCE FROM N.A.
RC      STRAIN=054G09;
RX      MEDLINE=20408577; PubMed=10952598;
RA      Bachelet L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA      Krakowski K., Bolling L., Aujay M., Wang X.V., Ellis D., Becker M.F.,
RA      Laust A.L., George H.J., Spalding D.R., Hollis G., Abremski K.;
RT      "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT      Failing Efavirenz Combination Therapy.";
RL      Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC      -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC      KNOWN AS THE RETROPEPSIN FAMILY.

EMBL: AY002119; AAF92274.1; -
DR      GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR      GO: GO:0008233; F:peptidase activity; IEA.

```

DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Asparticase AS.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept_A_acid.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00077; IVP; 1.
DR Pfam: PF00078; IVC; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
KM Aspartyl protease, Hydrolase; Protease; RNA-directed DNA polymerase;
Transferase.
KW NON TER 1 1
FT CHAIN 1 >99 PROTEASE.
FT CHAIN 100 >328 REVERSE TRANSCRIPTASE.
FT NON TER 328 328
SQ SEQUENCE 328 AA; 37138 MW; E7575063F438782F CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 17 RGRNPDIYQYWMDDL 32
||:|||||
Db 271 RGRNPDIYQYWMDDL 286

RESULT 15
Q9EPR2 PRELIMINARY; PRT; 328 AA.
AC Q9EPR2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=070605;
RX MEDLINE=20408577; PubMed=10952598;
RA Bachelier L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Boiling L., Anjay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abremski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy".
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL: AY002113; AAF92268.1; -.
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Asparticase AS.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept_A_acid.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00077; IVP; 1.
DR Pfam: PF00078; IVC; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
Transferase.
FT NON TER 1 1

FT CHAIN 1 >99 PROTEASE.
FT CHAIN 100 >328 REVERSE TRANSCRIPTASE.
FT NON TER 328 328
SQ SEQUENCE 328 AA; 37297 MW; A4704362BCB77CD CRC64;
Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 17 RGRNPDIYQYWMDDL 32
||:|||||
Db 271 RGRNPDIYQYWMDDL 286

Search completed: July 20, 2004, 06:33:18
Job time : 34.2623 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:07:17 ; Search time 47.7377 Seconds
(without alignments)
189,400 Million cell updates/sec

Title: US-09-673-166B-275
Perfect score: 168
Sequence: 1 GRQYIKANSKFIGITGRGNPDIVYQYMDL 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq 294n04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query length	DB ID	Description
1	98	58.3	28	2	AA53552
2	92.5	55.1	50	2	AAW06132
3	92.5	55.1	50	2	AAW46447
4	87	51.8	29	2	AA53551
5	86.5	51.5	25	4	AA884738
6	82.5	49.1	50	2	AAW06131
7	81	48.2	29	5	ABB79182
8	78.5	46.7	280	2	AAW71642
9	78.5	46.7	280	2	AAW71633
10	78.5	46.7	280	4	AAW71633
11	78	46.4	137	3	AAW82634
12	78	46.4	137	3	AAW82634
13	78	46.4	240	3	AAW49253
14	78	46.4	390	3	AAW49255
15	78	46.4	750	3	AAW92638
16	78	46.4	1003	3	AAW92638
17	78	46.4	1445	3	AAW92638
18	77	45.8	213	5	AAW92638
19	77	45.8	999	3	AAW92638
20	76.5	45.5	31	2	AAW06129
21	76.5	45.5	31	2	AAW06129
22	76.5	45.5	126	3	AAW45490
23	76	45.2	29	2	AAW83561
24	76	45.2	37	2	AAW5389
25	76	45.2	37	2	AAW5383

26	76	45.2	46	5	AAU11430	AAU11430 Synthetic
27	76	45.2	47	2	AAW62723	AAW62723 LHRH-cont
28	76	45.2	124	3	AAW45519	AAW45519 Modified
29	76	45.2	124	3	AAW45493	AAW45493 Modified
30	76	45.2	141	3	AAW45499	AAW45499 Modified
31	76	45.2	158	2	AAW81329	AAW81329 TNF2-4, a
32	76	45.2	373	4	AAW81048	AAW81048 Human TNF
33	76	45.2	750	3	AAW92627	AAW92627 Mutant hu
34	76	45.2	20	6	AAW53186	AAW53186 HIV he/pe
35	76	45.2	20	6	AAW53189	AAW53189 HIV CTL e
36	75	44.6	25	3	AAW92652	AAW92652 PSMep009
37	75	44.6	25	3	AAW92651	AAW92651 PSMep008
38	75	44.6	32	2	AAW26604	AAW26604 HIV-dariv
39	75	44.6	32	2	AAW83189	AAW83189 lipopepti
40	75	44.6	51	6	AAW35202	AAW35202 HIV epito
41	75	44.6	60	4	AAW94762	AAW94762 CTL epito
42	75	44.6	60	4	AAW89815	AAW89815 HIV Pol-2
43	75	44.6	60	5	AAW83192	AAW83192 HIV cytot
44	75	44.6	109	4	AAW20145	AAW20145 Growth di
45	75	44.6				

ALIGNMENTS

RESULT 1
AA53552
ID AA53552 standard; protein; 28 AA.

AC AA53552;
DT 18-JAN-2000 (first entry)
DE lipopeptide #3.

lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
human immunodeficiency virus; hepatitis B virus; papilloma virus;
melanoma; malaria; parasite.

OS Synthetic.
OS Homo sapiens.

Key Modified-site 1 Location/Qualifiers
FT 1
FT note="contains palmitoyl residue attached at the N-terminus and on the epsilon carbon of the side chain"

FR2776926-A1.

08-OCT-1999.

07-APR-1998; 98FR-00004323.

07-APR-1998; 98FR-00004323.

(INRM) INSERM INST NAT SANTE & RECH MEDICALE.

(CNRS) CNRS CENT NAT RECH SCI.

(INSP) INST PASTEUR LILLE.

Le Gal FA, Guillet JG, Gahery SH, Gras ME, Melnyk O, Tartar A;

WPI; 1999-583113/50.

New lipopeptide containing lipid regions and two epitopes, all separated by peptide spacers that impart hydrophilicity, useful in vaccines.

Example 1; Page 9; 35pp; French.

The invention relates to the generation of a lipopeptide comprising at least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL) epitope and at least one lipid residue with (i) the epitopes and lipid portion and (ii) the epitopes, being separated independently by peptide

CC spacers. These spacers comprise sequences of amino acids which carry an
CC overall electrical charge in neutral media to ensure that the lipopeptide
CC is hydrophilic. This peptide represents an example of them lipopeptide of
CC the invention. It contains 2 lipid residues attached at the N-terminal
CC end of the molecule and 2 epitopes separated by the spacer residues Ser-
CC Ser and Ala-Ala-Ala. Peptide AAY5301-Y5349 represent peptide epitopes
CC used in the generation of the lipopeptides. These are used in therapeutic
CC or prophylactic compositions and vaccines to induce specific immune
CC responses against human immunodeficiency, hepatitis B or papilloma
CC viruses; p53 of melanoma or the malaria parasite

SQ Sequence 26 AA;

Query Match 58.3%; Score 98; DB 2; Length 28;
Best Local Similarity 84.0%; Pred. No. 1e-07;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GROYIKANSKFIGITGRGRNPDIIV 25
Db 1 GROYIKANSKFIGITGRGRNPAGIGI 25

RESULT 2

ID AAM06132 standard; peptide; 50 AA.

AC AAM06132;

DT 07-FEB-1997 (first entry)

DE Anti-cholesterol ester transfer multivalent vaccine peptide.

KM Cholesterol ester transfer protein; CERP; antigen; vaccine;

KW cardiovascular disease; atherosclerosis.

XX Synthetic.

OS Location/Qualifiers

FH 2.15 /label= "T-cell epitope

FT /note= "T-cell epitope comprises amino acids 830-843 of

FT tetanus toxoid protein"

FT 16..34 /label= "B-cell epitope

FT /note= "B-cell epitope comprises amino acids 350-368 of

FT rabbit CERP"

FT 35..50 /label= "B-cell epitope

FT /note= "B-cell epitope comprises the C-terminal 16 amino

FT acids of rabbit CERP" involved in neutral lipid binding

FT or transfer activity"

XX Region

XX (TCEL-) T CELL SCI INC.

XX Rittershaus CM, Thomas LJ;

XX WPI; 1996-S06103/50.

XX Cholesterol ester transfer protein B cell epitope linked to T cell

XX epitope - used to generate vaccine to regulate CERP activity for

XX decreasing the risk of developing a cardiovascular disease e.g.

XX atherosclerosis.

XX Disclousure; Page 7; 72pp; English.

CC A multivalent vaccine comprises an immunogenic helper T-cell epitope of
CC tetanus toxoid protein covalently linked to the B-cell epitopes of rabbit
CC cholesterol ester transfer protein (CERP) (see also AAM06132). The
CC vaccine elicits an immune response against endogenous CERP activity, and
CC is used to treat or prevent a cardiovascular disease, such as
CC atherosclerosis

SQ Sequence 50 AA;

Query Match 55.1%; Score 92.5; DB 2; Length 50;
Best Local Similarity 60.6%; Pred. No. 1.4e-06;
Matches 20; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY 3 QYIKANSKFIGITGRNDP---IVYQYMDL 32
Db 2 QYIKANSKFIGITGRPPDGRBAVAYREEDI 34

RESULT 3

ID AAM46447 standard; peptide; 50 AA.

AC AAM46447;

DT 17-OCT-2003 (revised)

DT 27-AUG-2003 (revised)

DT 18-MAY-1998 (first entry)

DE CERP B cell epitope/tetanus toxoid construct for a plasmid vaccine.

KM Cholesterol ester transfer protein; CERP; cholesterol ester;

KW high density lipoprotein; HDL; very low density lipoprotein; VLDL;

KW low density lipoprotein; LDL; atherosclerosis; neutral lipid binding;

KW transfer activity; immunogenic; B cell epitope; antibody;

KW DNA plasmid-based vaccine; broad range helper T cell epitope; treatment;

KW cardiovascular disease; chimeric.

XX Clostridium tetani.

OS Oryctolagus sp.

OS Chimeric.

FH 2.15 Location/Qualifiers

FT Peptide /note= "tetanus toxoid broad range helper T cell epitope"

FT 16..34 /note= "CERP epitope of amino acids 350-368 of AAM46445"

FT 35..50 /note= "CERP epitope of amino acids 481-496 of AAM46445"

FT Peptide /note= "CERP epitope of amino acids 481-496 of AAM46445"

XX Region

XX (TCEL-) T CELL SCI INC.

XX Thomas LJ;

XX WPI; 1997-549731/50.

XX DNA plasmid-based vaccine encodes CERP B cell and helper T cell

XX epitope(s) - used for elevating high density lipoprotein levels, and for

XX treating cardiovascular disease.

XX Claim 8; Page 22; 67pp; English.

XX The present sequence represents a construct containing a helper T cell

XX epitope from the toxoid tetanus protein, and 2 B cell epitopes of rabbit

CC mature cholesterol ester transfer protein (CETP). CETPs mediate the
 CC transfer of cholesterol esters from high density lipoprotein (HDL) to
 CC very low density lipoprotein (VLDL) and low density lipoprotein (LDL),
 CC and vice versa. An increased CETP activity produces an atherogenic
 CC lipoprotein profile and induces atherosclerosis. A 13 amino acid stretch
 CC in the rabbit CETP (AAW46445) (Phe483 to Leu495) is important for neutral
 CC lipid binding and transfer activity. This region has been shown to be
 CC immunogenic as a B cell epitope of CETP. A second B cell epitope is
 CC defined by Arg350 to Ile368. Antibodies to this second epitope would
 CC allow the formation of immune complexes involving CETP, and promote the
 CC removal of the complexed CETP. The present sequence is encoded by a DNA
 CC plasmid-based vaccine which comprises sequences encoding at least 1 B
 CC cell epitope of CETP linked in frame with at least one segment encoding a
 CC broad range helper T cell epitope. The vaccines can be used to elevate
 CC the ratio of circulating HDL to circulating LDL, VLDL or total
 CC cholesterol in a human. It can also be used for decreasing the level of
 CC endogenous CETP activity in a human. The vaccine can be used to produce
 CC anti-CETP antibodies in vivo and for treating cardiovascular disease.
 CC (Updated on 27-AUG-2003 to correct OS field.) (Updated on 17-OCT-2003 to
 CC standardise OS field)

CC Sequence 50 AA;

Query Match 55.1%; Score 92.5; DB 2; Length 50;

Best Local Similarity 60.6%; Pred. No. 1.4e-06; Mismatches 6; Indels 3; Gaps 1;

OY 3 QYKANSKFGITGERGRNDP---IYIYQYMDL 32
 DB 2 QYKANSKFGITGERFPRPDGREAVAYRFEEDI 34

RESULT 4

AAV53551

ID AAV53551 standard; protein; 29 AA.

AC AAV53551;

DT 18-JAN-2000 (first entry)

DB Lipopeptide #2.

KW Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
 KW electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
 KW human immunodeficiency virus; hepatitis B virus; papilloma virus;
 KW melanoma; malaria; parasite.

OS Synthetic.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Modified-site 1 /note= "contains palmitoyl residue attached at the N-terminus and on the epsilon carbon of the side chain"

FR2776926-A1.

PD 08-OCT-1999.

PF 07-APR-1998; 98FR-00004323.

PR 07-APR-1998; 98FR-00004323.

PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PA (CNRS) CNRS CENT NAT RECH SCI.

PA (INSP) INST PASTEUR LILLE.

PI Le Gal FA, Guillet JG, Gabery SH, Gras MH, Melnyk O, Tartar A;

XX WPI, 1999-583113/50.

PT New lipopeptide containing lipid regions and two epitopes, all separated by peptide spacers that impart hydrophilicity, useful in vaccines.

XX Example 1; Page 9; 35pp; French.

PS The invention relates to the generation of a lipopeptide comprising at
 CC least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)
 CC epitope and at least one lipid residue with (i) the epitopes and lipid
 CC portion and (ii) the epitopes, being separated independently by peptide
 CC spacers. These spacers comprise sequences of amino acids which carry an
 CC overall electrical charge in neutral media to ensure that the lipopeptide
 CC is hydrophilic. This peptide represents an example of them lipopeptide of
 CC the invention. It contains 2 lipid residues attached at the N-terminal
 CC end of the molecule and 2 epitopes separated by the spacer residues Ser-
 CC Ser and Ala-Ala-Ala. Peptides AAY53301-Y53549 represent peptide epitopes
 CC used in the generation of the lipopeptides. These are used in therapeutic
 CC or prophylactic compositions and vaccines to induce specific immune
 CC responses against human immunodeficiency, hepatitis B or papilloma
 CC viruses, p53 of melanoma or the malaria parasite

XX Sequence 29 AA;

Query Match 51.8%; Score 87; DB 2; Length 29;

Best Local Similarity 82.6%; Pred. No. 5.2e-06; Mismatches 4; Indels 0; Gaps 0;

OY 3 QYKANSKFGITGERGRNDIV 25
 DB 4 QYKANSKFGITGERGRAAGIGI 26

RESULT 5

AAB84738

ID AAB84738 standard; peptide; 25 AA.

AC AAB84738;

DT 17-SEP-2001 (first entry)

DB Amino acid sequence of lipopeptide MH2-K.

KW Lipid-tailed protein; mucosal membrane; immune system; lipoprotein;
 KW B cell response; T cell response; intranasal immunisation;
 KW sublingual immunisation.

OS Synthetic.

FT Key Location/Qualifiers

FT Modified-site 1 /note= "amidated residue with palmitoyl lipid residue with 16 carbon atoms attached via an epsilon bond"

WO200141797-A2.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000WO-EP012794.

PR 09-DEC-1999; 99US-0169952P.

PA (INSP) INST PASTEUR.

PA (UTRI-) UNIV LILLE 2.

PI Drulhe P, Gras-Masse H, Benmohamed L;

XX WPI, 2001-451597/48.

PT Inducing systemic immune response comprises mucosal administration of lipid-tailed proteins or lipopeptide in absence of adjuvants.

PS Claim 22; Page 26; 34pp; English.

CC The present sequence represents a lipid-tailed protein. The lipid moiety is able to cross the mucosal membrane, and deliver an antigen to the immune system. The lipoprotein is administered to a mucosal membrane to

CC induce an immune response. The lipoprotein induces a B cell and/or a T
CC cell response. The lipoprotein is used to provide intranasal or
CC sublingual immunization
XX
SQ Sequence 25 AA;

Query Match 51.5%; Score 86.5; DB 4; Length 25;
Best Local Similarity 95.0%; Pred. No. 5.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 GROYI-KANSKFIGITERGR 19
1 GROYIKKANSKFIGITERGR 20

Db

RESULT 6
AAW06131
ID AAW06131 standard; peptide; 50 AA.
XX
AC AAW06131;
XX
DT 07-FEB-1997 (first entry)
XX
DE Anti-cholesterol ester transfer multivalent vaccine peptide.
XX
KW Cholesteryl ester transfer protein; CERP; antigen; vaccine;
XX cardiovascular disease; atherosclerosis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /note= "C-terminal Cys residue is present for use in
FT linking the peptide to itself or other molecules"
FT Region 2. .15
FT /label= "T-cell epitope
FT /note= "T-cell epitope comprises amino acids 830-843 of
FT tetanus toxoid protein"
FT Region 16. .34
FT /label= "B-cell epitope
FT /note= "B-cell epitope comprises amino acids 349-367 of
FT human CERP"
FT Region 35. .50
FT /label= "B-cell epitope
FT /note= "B-cell epitope comprises the C-terminal 16 amino
FT acids of human CERP. Involved in neutral lipid binding or
FT transfer activity"
XX
PN WO9634888-A1.
XX
PD 07-NOV-1996.
XX
PF 01-MAY-1996; 96WO-US0006147.
XX
PR 01-MAY-1995; 95US-00432483.
XX
PA (TCEL-) T CELL SCI INC.
XX
PI Rittershaus CW, Thomas LJ;
XX
DR WPI; 1996-506103/50.
XX
PT Cholesteryl ester transfer protein B cell epitope linked to T cell
PT epitope - used to generate vaccine to regulate CERP activity for
PT decreasing the risk of developing a cardiovascular disease e.g.
PT atherosclerosis.
XX
PS Disclosure; Page 7; 72pp; English.

CC A multivalent vaccine comprises an immunogenic helper T-cell epitope of
CC tetanus toxoid protein covalently linked to the B-cell epitopes of human
CC cholesteryl ester transfer protein (CERP) (see also AAW06127). The
CC vaccine elicits an immune response against endogenous CERP activity, and

CC is used to treat or prevent a cardiovascular disease, such as
CC atherosclerosis
XX
SQ Sequence 50 AA;

Query Match 49.1%; Score 82.5; DB 2; Length 50;
Best Local Similarity 57.6%; Pred. No. 5e-05;
Matches 19; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

3 OYIKANSKFIGITERGRND---IVLYQYMDL 32
2 OYIKANSKFIGITELFPPRPOQSHVAYTEEDI 34

Db

RESULT 7
ABB79182
ID ABB79182 standard; peptide; 29 AA.
XX
AC ABB79182;
XX
DT 07-AUG-2002 (first entry)
XX
DE Human cytomegalovirus PPI50 related vaccine peptide SEQ ID NO:8.
XX
KW Human cytomegalovirus PPI50; HCMV; human leukocyte antigen; HLA; CTL;
KW cytotoxic T lymphocyte; immunoreactive; epitope; vaccine; infection;
KW immunogenic; virucide; immune response.
XX
OS Human cytomegalovirus.
XX
PN WO200234769-A2.
XX
PD 02-MAY-2002.
XX
PF 22-OCT-2001; 2001WO-US032589.
XX
PR 20-OCT-2000; 2000US-0241944P.
XX
PA (CITY) CITY OF HOPE.
XX
PI Diamond DJ;
XX
DR WPI; 2002-471432/50.
XX
PT New immunoreactive peptide cytotoxic T-lymphocytes epitope of human
PT cytomegalovirus PPI50, useful for creating vaccines against
PT cytomegalovirus infection.
XX
PS Disclosure; Page 14; 28pp; English.

CC The present invention describes a peptide (I) which is an immunogenic
CC epitope recognised by CD8+ class I major histocompatibility complex (MHC)
CC restricted cytotoxic T-lymphocytes (CTL) of patients harbouring latent
CC cytomegalovirus (HCMV) infection. Also described are: (1) a vaccine (II)
CC against human cytomegalovirus (CMV), which comprises (1); (2) a cellular
CC vaccine (III) against human CMV comprises antigen presenting cells which
CC present (1); (3) a recombinant viral vector (IV) which contains a gene
CC encoding (1); and (4) an immunological reagent (V) which comprises (1).
CC (1) has virucide activity. (II) or (III) are useful for modulating the
CC immune response to HCMV infection, and for vaccinating a mammal in need
CC of it against HCMV. (1) is useful for eliciting a cellular immune against
CC HCMV by normal and immunodeficient subjects, for creating efficient
CC vaccines against HCMV, and in immunological methods to detect PPI50-
CC reactive CTL in a patient or a sample from the patient. The present
CC sequence represents a HCMV PPI50 related vaccine peptide, which is given
CC in the exemplification of the present invention
XX
SQ Sequence 29 AA;

Query Match 48.2%; Score 81; DB 5; Length 29;
Best Local Similarity 100.0%; Pred. No. 4.3e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GROUYKANSKFIGITE 16
 DB 14 GROUYKANSKFIGITE 29

RESULT 8
 AAW71642
 ID AAW71642 standard; protein; 280 AA.
 XX
 AC AAW71642;
 XX
 DT 17-OCT-2003 (revised)
 DT 18-JAN-1999 (first entry)
 XX
 DE HIV POL/NEF epitopes.
 XX
 KM Vector; gene therapy; vaccine; ALVAC; translation factor; K3L; E3L;
 KM VCP1433; HIV; pol; nef; epitope.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN WO9840500-A1.
 XX
 PD 17-SEP-1998.
 XX
 PF 25-FEB-1998; 98MO-US003710.
 XX
 PR 12-MAR-1997; 97US-00815809.
 XX
 PA (VIR-) VIROGENETICS CORP.
 PA (UYAR-) UNIV ARIZONA STATE.
 XX
 PI Tartaglia J, Jacobs BL, Goebel SJ, Cox WI, Gettig RR, Pincus SE;
 PI Paoletti E;
 XX
 DR WPI; 1998-520819/44.
 DR N-PSDB; AAV58243.
 XX

Enhancing expression of nucleic acids in cells - by using modified
 PT vectors which comprise the nucleic acid and also nucleic acid encoding a
 PT translation factor.
 XX

Example 1; Fig 4A-C; 90pp; English.

This polypeptide comprises the POL and NEF epitopes of HIV-1. It is
 encoded by VCP1433 (see AAV58243). In VCP1433, an HIV pol/nef 'string of
 beads' cassette is placed under control of the vaccinia H6 promoter.
 CC pMC6H6K33, containing a vaccinia H6/K3L expression cassette and
 CC vaccinia E3L gene with endogenous promoter flanked by the ALVAC C6
 CC insertion site sequences, was used in recombination with VCP1433 to
 CC obtain VCP1452 (see AAV58244-45). K3L and E3L are vaccinia virus
 CC translation factors. Novel vectors of the invention, such as ALVAC
 CC vectors, include K3L and/or E3L and are used for enhancing expression of
 CC gene products that they encode. The translation factors can effect
 CC inhibition of eIF-2alpha phosphorylation or inhibition of protein kinase
 CC PKR phosphorylation or otherwise sequester double stranded (ds) RNA,
 CC increasing the effective concentration of ds RNA. The up-regulation of a
 CC foreign gene expression can have a profound effect on the induction of a
 CC therapeutic or immunological response in a host administered or
 CC inoculated with recombinants derived from these new vectors. (Updated on
 CC 17-OCT-2003 to standardise OS field)
 CC
 XX

Sequence 280 AA;

Query Match 46.7%; Score 78.5; DB 2; Length 280;
 Best Local Similarity 47.5%; Pred. No. 0.0017;
 Matches 19; Conservative 4; Mismatches 6; Indels 11; Gaps 2;

QY 3 QYIKANSKFIGITE-----RGRNPDIYIYQWDDL 32
 DB 81 EYFK-NCKLMAIFQSSMTKILPEPRKONPDIVIYQWDDL 119

RESULT 9
 AAW71363
 ID AAW71363 standard; protein; 280 AA.
 XX
 AC AAW71363;
 XX
 DT 17-OCT-2003 (revised)
 DT 15-FEB-1999 (first entry)
 XX
 DE HIV POL/NEF epitopes.
 XX
 KM Vector; gene therapy; vaccine; ALVAC; K3L; E3L; translation factor;
 KM VCP1433; HIV; POL; NEF; epitope.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN WO9840501-A1.
 XX
 PD 17-SEP-1998.
 XX
 PF 13-FEB-1998; 98MO-US002669.
 XX
 PR 12-MAR-1997; 97US-00816155.
 XX
 PA (VIR-) VIROGENETICS CORP.
 XX
 PI Tartaglia J, Cox WI, Gettig RR, Martinez H, Paoletti E;
 PI Pincus SE;
 XX
 DR WPI; 1998-520820/44.
 DR N-PSDB; AAV60251.
 XX

Enhancing expression of nucleic acids in cells - by using modified
 PT vectors which comprise nucleic acid and also nucleic acid encoding
 PT transcription factor and optionally translation factor.
 XX

Example 2; Fig 6; 102pp; English.

This polypeptide comprises the POL and NEF epitopes of HIV-1. It is
 encoded by VCP1433 (see AAV60251). In VCP1433, an HIV pol/nef 'string of
 CC beads' cassette is placed under control of the vaccinia H6 promoter.
 CC pMC6H6K33, containing a vaccinia H6/K3L expression cassette and
 CC vaccinia E3L gene with endogenous promoter flanked by the ALVAC C6
 CC insertion site sequences, was used in recombination with VCP1433 to
 CC obtain VCP1452 (see AAV60252-53). K3L and E3L are vaccinia virus
 CC translation factors. New vectors are provided for enhanced expression of
 CC at least 1 first nucleic acid molecule (NAM) in a cell having a
 CC particular phenotype. The vector (e.g. NYVAC or ALVAC) is modified to
 CC comprise the first NAM and at least 1 second NAM encoding a transcription
 CC factor (TF), or a TF and a translation factor such as K3L and E3L, where
 CC there is co-temporal expression of the first and second NAMS with respect
 CC to the phenotype of the cell, and where expression of the second NAM
 CC enhances expression of the first NAM by enhancing transcription or
 CC transcription and translation. Also claimed is a method for increasing
 CC expression of at least 1 first NAM by such a vector. The vectors can be
 CC used for increasing expression of e.g. an epitope of interest, a
 CC biological response modulator, a growth factor, a recognition sequence, a
 CC therapeutic gene or a fusion protein. (Updated on 17-OCT-2003 to
 CC standardise OS field)
 CC
 XX

Sequence 280 AA;

Query Match 46.7%; Score 78.5; DB 2; Length 280;
 Best Local Similarity 47.5%; Pred. No. 0.0017;
 Matches 19; Conservative 4; Mismatches 6; Indels 11; Gaps 2;

QY 3 QYIKANSKFIGITE-----RGRNPDIYIYQWDDL 32
 DB 81 EYFK-NCKLMAIFQSSMTKILPEPRKONPDIVIYQWDDL 119

RESULT 10
 AAB98912

XX	AA	B98912	standard; protein, 280 AA.
AC	AA	B98912;	
XX	DT	25-SEP-2001	(first entry)
XX	DE	Synthetic HIV-derived Pol-Nef chimeric protein.	
XX	KM	HIV-specific immunogen; major histocompatibility complex; MHC class I;	
KM	KM	MHC class II; antigen presentation; CD8+ response; CD4+ response;	
KW	KW	gene therapy; genetic vaccine; anti-HIV; antiviral; antiretroviral; Pol;	
KW	XX	Nef; chimeric protein; CTL epitope; cytotoxic T-lymphocyte.	
XX	OS	Synthetic.	
XX	OS	Human immunodeficiency virus.	
FH	FH	Key	Location/Qualifiers
FT	FT	Region	1..60
FT	FT	Region	/note="Pol-3 CTL epitope (AAB98913)"
FT	FT	Region	61..86
FT	FT	Region	/note="Nef-2 CTL epitope (AAB98914)"
FT	FT	Region	89..148
FT	FT	Region	/note="Pol-2 CTL epitope (AAB98915)"
FT	FT	Region	149..231
FT	FT	Region	/note="Nef-1 CTL epitope (AAB98916)"
FT	FT	Region	232..280
FT	FT	Region	/note="Pol-1 CTL epitope (AAB98917)"
FN	FN	WO200154701-A1.	
XX	PD	02-AUG-2001.	
XX	PF	26-JAN-2001; 2001WO-US002766.	
XX	PR	31-JAN-2000; 2000US-0179276P.	
XX	PA	(AVENTIS PASTEUR SA.	
XX	PA	(AARO-) AARON DIAMOND AIDS RES CENT.	
XX	PI	Ho D, Markowitz M, Klein M, El Habib R;	
XX	PI	WPI, 2001-488741/53.	
XX	PT	Permitting cessation of antiviral therapy on HIV-infected patients	
XX	PT	undergoing antiviral therapy, useful for treating HIV-infected patients,	
XX	PT	by administering nucleic acid based vaccines encoding HIV-specific	
XX	PT	immunogens.	
PS	PS	Disclosure; Page 12; 57pp; English.	
XX	CC	The invention relates to a method for permitting the cessation of	
XX	CC	antiviral therapy in HIV-infected patients undergoing antiviral therapy	
XX	CC	who have a controlled level of viraemia. The method involves	
XX	CC	administration of nucleic-acid based vaccines encoding HIV-specific	
XX	CC	immunogens to the patient, enabling the expression of the HIV-specific	
XX	CC	immunogens in the patient's cells. The HIV-specific antigens are	
XX	CC	presented on major histocompatibility complex (MHC) class I and II	
XX	CC	molecules, stimulating an HIV-specific CD8+ and CD4+ response. The method	
XX	CC	permits the cessation of antiviral therapy on HIV-infected patients	
XX	CC	either without virus rebound, with a delayed rebound, or with decreased	
XX	CC	post-rebound set point, in addition to achieving immunological control of	
XX	CC	persistent infectious virus after discontinuation of antiviral therapy.	
XX	CC	The method is useful for people who have lost their CD4+ and/or CD8+ T	
XX	CC	cell responses to HIV antigens, such as those who have lost their	
XX	CC	proliferative T cell responses to gp120 or p24. The method is	
XX	CC	specifically useful for treating persons infected with a lymphotropic or	
XX	CC	immune-destroying retroviral infection, or for treating HIV-infected	
XX	CC	patients undergoing anti-retroviral therapy and having a viral load of	
XX	CC	less than 10000, preferably 1000 viral copies/ml of plasma and a CD4+	
XX	CC	cell count of above 300 cells/ml, preferably 500 cells/ml. The present	
XX	CC	sequence represents a synthetic protein comprising N- to C-terminally,	
XX	CC	the HIV-specific cytotoxic T-lymphocyte (CTL) epitopes Pol-3 (AAB98913),	
XX	CC	Nef-2 (AAB98914), Pol-2 (AAB98915), Nef-1 (AAB98916) and Pol-1 (AAB98917)	

SX Sequence 280 AA; 46.7%; Score 78.5; DB 4; Length 280;
Query Match Best Local Similarity 47.5%; Pred. No. 0.0017;
Matches 19; Conservative 4; Mismatches 6; Indels 11; Gaps 2

CY 3 QYIKANSKEFIGITE-----RGRNPDPVITYQYMDL 32
| | | | |
DB 81 EYFK-NCTLMALFQSSWTKILEPPRKPNPDIVITYQYMDL 119

RESULT 11
ID AAY82634
AAY82634 standard; peptide; 137 AA.

AC AAY82634;
DT 07-AUG-2000 (First entry)
DE Tetanus toxoid T cell epitopes and Der pII B cell epitopes peptide.
XX
XX T cell epitope; B cell epitope; allergy; allergen; antigenic;
KM antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
KM atopic dermatitis; acute urticaria; chronic urticaria;
KW gastro-intestinal syndrome; food allergy; oro-pharyngeal syndrome;
KM anaphylactic reaction; drug hypersensitivity; allergic reaction.
XX
XX Dermatophagoides pteronyssinus.
OS Clostridium tetani.
OS Synthetic.
XX
XX WO200006694-A2.
XX
XX PD 10-FEB-2000.
XX PF 20-JUL-1999; 99WO-BE000092.
XX PR 30-JUL-1998; 98EP-00870167.
XX PA (UNIO) UCB SA.
XX PI Saint-Remy J, Jacquemin M;
XX WPI; 2000-422470/36.
XX
XX New compound for prevention and treatment of allergies comprises at least one allergen antigenic determinant recognized by a B cell and at least one antigenic determinant which does not trigger T cell activation.
XX
XX Claim 8; Page 35; 50pp; English.
XX
XX The present invention describes a compound (I) for the prevention and/or treatment of allergy. The compound comprises at least one allergen antigenic determinant (i) recognised by a B cell or an antibody secreted by a B cell of a non-atopic individual and at least one antigenic determinant (ii) different from the allergen that triggers T cell activation. (I) has antiallergic, antiasthmatic, antiinflammatory, dermatological and immunosuppressive activities, and can be used in a vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to treat and/or prevent allergies or a disease of allergic origin, especially hypersensitivities. These include rhinitis, sinusitis, bronchial asthma, atopic dermatitis, some forms of acute and chronic urticaria, gastro-intestinal syndromes associated with the ingestion of food allergens, oro-pharyngeal syndrome, anaphylactic reactions associated with drug hypersensitivities and/or a mixture of these. The use of (I) in the treatment of allergic conditions avoids the need for drug treatment, which often causes undesirable side-effects. Also, prior art drug therapies alleviate symptoms, but do not influence their causes, however (I) actually combats the cause of an allergic reaction. The present sequence represents a specifically claimed compound peptide sequence from the present invention

```
XX
SQ Sequence 137 AA;
Query Match 46.4%; Score 78; DB 3; Length 137;
Best Local Similarity 64.3%; Pred. No. 0.0085;
Matches 18; Conservative 1; Mismatches 5; Indels 4; Gaps 1;
QY 1 GROYIKANSKFIGITE---RGRNPDI 24
Db 17 GGOYIKANSKFIGITELSSCHSGSPCII 44
RESULT 12
AAV49253
ID AAV49253 standard; protein; 218 AA.
AC AAV49253;
XX
XX 07-FEB-2000 (first entry)
XX
DE N10 polypeptide carrier protein construct amino acid sequence.
XX
XX Carrier protein; CD4+; T cell epitope; N6; N10; N19; immunogen;
XX encapsulated bacteria.
XX
XX Synthetic.
XX
XX WO9955730-A2.
XX
XX 04-NOV-1999.
XX
XX 27-APR-1999; 99WO-IB000844.
XX
XX 27-APR-1998; 98GB-00008932.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Rappuoli R, Grandi G;
XX
XX WPI; 2000-023325/02.
XX
XX N-PSDB; AAZ31415.
XX
XX Carrier proteins containing CD4+ epitopes useful for protecting against
XX diseases caused by encapsulated bacteria.
XX
XX Disclosure; Fig 2; 76pp; English.
XX
XX The invention provides carrier proteins comprising at least 5 CD4+ T cell
XX epitope. The carrier protein comprises at least 1 of N6, N10 or N19. The
XX carrier protein can be prepared by expressing a vector comprising a
XX nucleic acid molecule encoding the protein, in a host cell and recovering
XX the expressed protein. The carrier protein can also be produced by (a)
XX constructing oligonucleotide molecules that encode peptide epitopes; (b)
XX annealing the oligonucleotides to form duplexes; (c) introducing the
XX duplexes into an expression vector; (d) introducing the expression vector
XX into a host cell; and (e) isolating the fusion protein produced from a
XX culture of the host cells. The carrier protein can be used as a
XX protective immunogen in the control of diseases caused by encapsulated
XX bacteria. The present sequence represents the amino acid sequence of N10
XX polypeptide carrier protein construct
XX
XX Sequence 218 AA;
Query Match 46.4%; Score 78; DB 3; Length 218;
Best Local Similarity 78.9%; Pred. No. 0.0015;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 3 QYIKANSKFIGITERGNP 21
Db 126 QYIKANSKFIGITERKGGSP 144
RESULT 13
```

```
AAV49254
ID AAV49254 standard; protein; 240 AA.
XX
XX AAV49254;
XX
XX 07-FEB-2000 (first entry)
XX
DE N11 polypeptide carrier protein construct amino acid sequence.
XX
XX Carrier protein; CD4+; T cell epitope; N6; N10; N19; immunogen; N11;
XX encapsulated bacteria.
XX
XX Synthetic.
XX
XX WO9955730-A2.
XX
XX 04-NOV-1999.
XX
XX 27-APR-1999; 99WO-IB000844.
XX
XX 27-APR-1998; 98GB-00008932.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Rappuoli R, Grandi G;
XX
XX WPI; 2000-023325/02.
XX
XX N-PSDB; AAZ31416.
XX
XX Carrier proteins containing CD4+ epitopes useful for protecting against
XX diseases caused by encapsulated bacteria.
XX
XX Disclosure; Fig 7; 76pp; English.
XX
XX The invention provides carrier proteins comprising at least 5 CD4+ T cell
XX epitope. The carrier protein comprises at least 1 of N6, N10 or N19. The
XX carrier protein can be prepared by expressing a vector comprising a
XX nucleic acid molecule encoding the protein, in a host cell and recovering
XX the expressed protein. The carrier protein can also be produced by (a)
XX constructing oligonucleotide molecules that encode peptide epitopes; (b)
XX annealing the oligonucleotides to form duplexes; (c) introducing the
XX duplexes into an expression vector; (d) introducing the expression vector
XX into a host cell; and (e) isolating the fusion protein produced from a
XX culture of the host cells. The carrier protein can be used as a
XX protective immunogen in the control of diseases caused by encapsulated
XX bacteria. The present sequence represents the amino acid sequence of N11
XX polypeptide carrier protein construct
XX
XX Sequence 240 AA;
Query Match 46.4%; Score 78; DB 3; Length 240;
Best Local Similarity 78.9%; Pred. No. 0.0017;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 3 QYIKANSKFIGITERGNP 21
Db 126 QYIKANSKFIGITERKGGSP 144
RESULT 14
AAV49255
ID AAV49255 standard; protein; 390 AA.
XX
XX AAV49255;
XX
XX 07-FEB-2000 (first entry)
XX
DE N19 polypeptide carrier protein construct amino acid sequence.
XX
XX Carrier protein; CD4+; T cell epitope; N6; N10; N19; immunogen;
XX encapsulated bacteria.
XX
XX Synthetic.
XX
```

```
XX XX WO9955730-A2.
XX XX 04-NOV-1999.
XX XX 27-APR-1999; 99WO-IB000844.
XX XX 27-APR-1998; 98GB-00008932.
XX XX (CHIR-) CHIRON SPA.
XX XX Rappuoli R, Grandi G;
XX XX WPI: 2000-023325/02.
XX XX N-PSDB; AAZ31417.
XX XX Carrier proteins containing CD4+ epitopes useful for protecting against
XX XX diseases caused by encapsulated bacteria.
XX XX Disclosure; Fig 8; 76pp; English.
XX XX The invention provides carrier proteins comprising at least 5 CD4+ T cell
XX XX epitope. The carrier protein comprises at least 1 of N6, N10 or N19. The
XX XX carrier protein can be prepared by expressing a vector comprising a
XX XX nucleic acid molecule encoding the protein, in a host cell and recovering
XX XX the expressed protein. The carrier protein can also be produced by (a)
XX XX constructing oligonucleotide molecules that encode peptide epitopes; (b)
XX XX annealing the oligonucleotides to form duplexes; (c) introducing the
XX XX duplexes into an expression vector; (d) introducing the expression vector
XX XX into a host cell; and (e) isolating the fusion protein produced from a
XX XX culture of the host cells. The carrier protein can be used as a
XX XX protective immunogen in the control of diseases caused by encapsulated
XX XX bacteria. The present sequence represents the amino acid sequence of N19
XX XX polypeptide carrier protein construct
XX XX
XX XX Sequence 390 AA;
XX XX
XX XX Query Match 46.4%; Score 78; DB 3; Length 390;
XX XX Best Local Similarity 78.9%; Pred. No. 0.0031; 2; Indels 0; Gaps 0;
XX XX Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX XX
XX XX 3 QYIKANSKFIGITERGRNP 21
XX XX ||||| ||||| ||||| :||
XX XX Db 126 QYIKANSKFIGITERKGGSP 144
XX XX
XX XX RESULT 15
XX XX ID AAY92638
XX XX AAY92638 standard; protein; 750 AA.
XX XX
XX XX AAY92638;
XX XX
XX XX 10-AUG-2000 (first entry)
XX XX
XX XX Mutant human prostate specific membrane antigen construct, hPSM3.1.
XX XX
XX XX Prostate specific membrane antigen; immunogenized construct; mutant;
XX XX KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
XX XX prostate cancer; cell-associated peptide antigen; foreign epitope.
XX XX
XX XX Homo sapiens.
XX XX OS Synthetic.
XX XX
XX XX Key Location/Qualifiers
XX XX FH 21..41
XX XX FT Peptide
XX XX /label= P30
XX XX /note= "Foreign epitope"
XX XX FT 213..227
XX XX /label= P2
XX XX /note= "foreign epitope"
XX XX
XX XX WO200020027-A2.
XX XX
```

```
PD 13-APR-2000.
XX XX
XX XX 05-OCT-1999; 99WO-DK000525.
XX XX
XX XX 05-OCT-1998; 98DK-00001261.
XX XX PR 20-OCT-1998; 98US-0105011P.
XX XX
XX XX (MERI-) M & B BIOTECH AS.
XX XX
XX XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX XX PI Gautam A, Birk P, Karlsson G;
XX XX WPI: 2000-349917/30.
XX XX
XX XX Inducing immune responses to weakly immunogenic, tumor associated peptide
XX XX PT antigens for the treatment of breast and prostate cancer.
XX XX
XX XX Example 1; Page; 220pp; English.
XX XX
XX XX AAY92627-49 are mutant immunogenized human prostate specific membrane
XX XX antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
XX XX The immunogenic analogues of PSM can be used in the claimed method as an
XX XX autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
XX XX binding regions and cysteine residues involved in disulfide bonds are
XX XX preserved in the immunogenized forms. The method is used for inducing
XX XX immune responses against weakly immunogenic cell-associated peptide
XX XX antigens (PA) such as those associated with cancers (self-proteins), e.g.
XX XX human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
XX XX fibroblast growth factor 8b (FGF8b). The method comprises effecting
XX XX CC simultaneous presentation by antigen producing cells (APCs) of the
XX XX CC animal's immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
XX XX CC group derived from the PA and/or at least 1 B-cell group derived from the
XX XX CC cell-associated PA; and (2) at least 1 first T helper cell group which is
XX XX CC foreign to the animal. Analogues of human PSM, human Her2 and
XX XX CC human/murine FGF8b comprising a substantial part of all known and
XX XX CC predicted CTL and B-cell epitopes of the respective PA and including at
XX XX CC least one foreign T helper epitope are also claimed. The method is used
XX XX CC to treat prostate, prostate/breast or breast cancer when the PA is human
XX XX CC PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
XX XX CC from the wild type human PSM (AAY92619), which appears on pages 184-187
XX XX CC of the specification
XX XX
XX XX Sequence 750 AA;
XX XX
XX XX Query Match 46.4%; Score 78; DB 3; Length 750;
XX XX Best Local Similarity 76.2%; Pred. No. 0.0069; 4; Indels 0; Gaps 0;
XX XX Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX XX
XX XX 1 GRQYIKANSKFIGITERGRNP 21
XX XX | ||||| ||||| ||||| :||
XX XX Db 211 GRQYIKANSKFIGITELYSDP 231
XX XX
```

Search completed: July 20, 2004, 06:30:28
Job time : 48.7377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:10:18 ; Search time 11.7049 Seconds

(without alignments)
279.413 Million cell updates/sec

Title: US-09-673-166b-276

Perfect score: 180

Sequence: 1 GRQYIKANSKFRIGTERGRFPVTPQVPLRPMTYK 34

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR78:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	85	47.2	206	1 ASLJ12	nef protein - huma
2	85	47.2	206	1 ASLJ12	nef protein - huma
3	85	47.2	206	2 S03245	nef protein (clone
4	83	46.1	204	2 S24985	nef protein - huma
5	83	46.1	205	1 ASLJH3	nef protein - huma
6	83	46.1	205	2 S43467	nef protein - huma
7	83	46.1	205	2 JC7975	HIV accessory prot
8	83	46.1	206	1 ASLJFV	nef protein (clone
9	83	46.1	206	2 S03244	nef protein - huma
10	83	46.1	206	2 S33986	nef protein (clone
11	83	46.1	206	2 JC5400	nef protein - huma
12	79	43.9	206	2 S25937	nef protein - huma
13	78	43.3	97	2 S54385	nef protein - huma
14	78	43.3	207	1 QQLJND	nef protein - huma
15	78	43.3	212	1 QQLJZR	nef protein - huma
16	77	42.8	182	2 S03247	nef protein (clone
17	77	42.8	209	2 T01673	nef protein - huma
18	77	42.8	210	1 ASLJ02	nef protein - huma
19	77	42.8	218	1 ASLJBR	nef protein - huma
20	74	41.1	205	1 B44963	nef protein - huma
21	73	40.6	204	2 S03246	nef protein (clone
22	72	40.0	205	1 ASLJIK	nef protein - huma
23	72	40.0	214	1 I44001	nef protein - huma
24	70	38.9	1315	1 BICLJN	nef protein - huma
25	65	36.1	229	1 ASLJGN	nef protein - huma
26	65	36.1	255	1 ASLJGN	nef protein - huma
27	62	34.4	255	1 ASLJG2	nef protein - huma
28	61	33.9	250	2 S54851	nef protein - huma
29	61	33.9	256	1 ASLJH2	nef protein - huma

30	61	33.9	260	1 ASLJG2	nef protein - huma
31	60	33.3	240	2 S24572	nef protein - huma
32	59	32.8	257	2 S12160	nef protein - huma
33	58	32.2	226	2 S46353	nef protein - simi
34	56	31.1	251	2 S54867	nef protein - simi
35	56	31.1	255	2 S53099	nef protein - huma
36	55	30.6	238	2 S61205	nef protein - simi
37	55	30.6	239	2 S54852	nef protein - simi
38	55	30.6	244	2 S61208	nef protein - simi
39	55	30.6	246	2 S54853	nef protein - simi
40	55	30.6	250	2 S54849	nef protein - simi
41	55	30.6	250	2 S61207	nef protein - simi
42	54.5	30.3	3430	1 GNMWV	genome polypeptide
43	52	28.9	209	2 S61204	nef protein - simi
44	52	28.9	211	1 ASLJMS	nef protein - simi
45	52	28.9	239	2 S61206	nef protein - simi

ALIGNMENTS

RESULT 1

ASLJ12
nef protein - human immunodeficiency virus type 1 (isolate HTLV-III, 12)

M/Alternate names: 3'-orf protein; orf-F protein

C/Species: human immunodeficiency virus type 1, HIV-1

A/Note: host Homo sapiens (man)

C/Date: 04-Dec-1986 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999

C/Accession: A04006

R/ARYA, S.K.; Gallo, R.C.
Proc. Natl. Acad. Sci. U.S.A. 83, 2209-2213, 1986

A/Title: Three novel genes of human T-lymphotropic virus type III: immune reactivity of

A/Reference number: A94093; MUID:8617573; PMID:3008154

A/Accession: A04006

A/Molecule type: DNA

A/Residues: 1-206 <ARY>

A/Cross-References: EMBL:M11840; NID:G328453; PIDN:AAA45001.1; PID:G328458

A/Genetics:

A/Gene: nef; 3'-orf; orf-F

C/Superfamily: AIDS nef protein

C/Keywords: AIDS; immunodeficiency

Query Match

Best Local Similarity 47.2%; Score 85; DB 1; Length 206;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRPPVTPQVPLRPMTYK 34

DB 64 ERGRPPVTPQVPLRPMTYK 82

RESULT 2
nef protein - human immunodeficiency virus type 1 (isolate LV)

M/Alternate names: 3'-orf protein; orf-F protein

C/Species: human immunodeficiency virus type 1, HIV-1

C/Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999

C/Accession: A04007

R/Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Laszky, L.A.; Capon, D.J.

A/Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovi

A/Reference number: A93355; MUID:8511157; PMID:2982104

A/Accession: A04007

A/Molecule type: DNA

A/Residues: 1-206 <MUE>

A/Cross-References: GB:X02083; NID:G555008; PIDN:AA559874.1; PID:G328560

A/Genetics:

A/Gene: nef; 3'-orf; orf-F

C/Superfamily: AIDS nef protein

C/Keywords: AIDS; immunodeficiency

Query Match

Best Local Similarity 47.2%; Score 85; DB 1; Length 206;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMTYK 34

Db 64 EKVGFPVTPQVPLRPMTYK 82

RESULT 3
nef protein (clone HXB3) - human immunodeficiency virus type 1

N/Alternate names: 3'-orf protein

C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 20-Sep-1999

C/Accession: S03245

R/Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, R.P.; Livak, K.J.; Pettena
Nucleic Acids Res. 13, 8219-8229, 1985

A/Title: Polymorphism of the 3' open reading frame of the virus associated with the acqy
A/Reference number: S03244; PMID:86067228; PMID:2999715

A/Accession: S03245

A/Molecule type: DNA

A/Residues: 1-206 <RAT>
A/Cross-references: EMBL:X03198; NID:g61556; PIDN:CAA26947.1; PID:g61557

C/Genetics:

A/Gene: nef; 3'-orf; orf-F
C/Superfamily: AIDS nef protein

Query Match 47.2%; Score 85; DB 2; Length 206;

Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMTYK 34

Db 64 EKVGFPVTPQVPLRPMTYK 82

RESULT 4

nef protein - human immunodeficiency virus type 1

C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999

C/Accession: S24985

R/Harris, M.; Hsiolop, S.; Patsillacos, P.; Neil, J.C.
submitted to the EMBL Data Library, November 1991

A/Description: In vivo derived HIV-1 NEF genes are heterogeneous and lack detectable nuc
A/Reference number: S24985

A/Accession: S24985

A/Status: preliminary
A/Molecule type: DNA

A/Residues: 1-204 <HAR>
A/Cross-references: EMBL:X63044; NID:g60123; PIDN:CAA44770.1; PID:g60124

C/Superfamily: AIDS nef protein

Query Match 46.1%; Score 83; DB 2; Length 204;

Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMTYK 34

Db 66 FPVTPQVPLRPMTYK 80

RESULT 5

nef protein - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)

N/Alternate names: 3'-orf protein; orf-F protein

C/Species: human immunodeficiency virus type 1, HIV-1
A/Note: host Homo sapiens (man)
C/Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 31-Jan-1997

C/Accession: A04005

R/Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Dor
nberger, J.A.; Papas, T.S.; Garbayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.
Nature 313, 277-284, 1985

A/Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.

A/Reference number: A93353; PMID:85111123; PMID:2578615

A/Accession: A04005
A/Molecule type: DNA

A/Residues: 1-205 <RAT>
C/Genetics:

A/Gene: nef; 3'-orf; orf-F
C/Superfamily: AIDS nef protein
C/Keywords: AIDS; immunodeficiency

Query Match 46.1%; Score 83; DB 1; Length 205;

Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMTYK 34

Db 68 FPVTPQVPLRPMTYK 82

RESULT 6

nef protein - human immunodeficiency virus type 1

C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 20-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 03-May-1996

C/Accession: S43467

R/Freund, J.; Kellner, R.; Houthaeve, T.; Kalbitzer, H.R.
Eur. J. Biochem. 221, 811-819, 1994

A/Title: Stability and proteolytic domains of Nef protein from human immunodeficiency vi
A/Reference number: S43467; PMID:94229079; PMID:8174561

A/Accession: S43467

A/Status: preliminary
A/Molecule type: protein

A/Residues: 1-205 <FRE>
C/Superfamily: AIDS nef protein

Query Match 46.1%; Score 83; DB 2; Length 205;

Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMTYK 34

Db 67 FPVTPQVPLRPMTYK 81

RESULT 7

HIV accessory protein Nef - mouse

C/Species: Mus musculus (house mouse)
C/Date: 25-Aug-2003 #sequence_revision 25-Aug-2003 #text_change 15-Sep-2003

C/Accession: J07975

R/Sacco, P.; Corella, D.; Santoro, C.
Biochem. Biophys. Res. Commun. 305, 1061-1066, 2003

A/Title: Selection of peptides with affinity for the N-terminal domain of GATRA-1: Identiti
A/Reference number: J07975; PMID:12767938

A/Accession: J07975
A/Molecule type: DNA

A/Residues: 1-205 <SEC>
C/Comment: This protein, an accessory viral protein, acts as a nuclear regulatory factor
binds p53 via its amino-terminal domain and protects cell against p53-mediated apoptotic

C/Keywords: GATRA-1; Nef

Query Match 46.1%; Score 83; DB 2; Length 205;

Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMTYK 34

Db 68 FPVTPQVPLRPMTYK 82

RESULT 8

nef protein - human immunodeficiency virus type 1 (isolate LAV-1a)

N/Alternate names: 3'-orf protein; orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
 C:Accession: A04008; S14609
 R:Main-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.
 Cell 40, 9-17, 1985
 A:Title: Nucleotide sequence of the AIDS virus, LAV.
 A:Reference number: A90866; MUID:8509333; PMID:2981635
 A:Accession: A04008
 A:Molecule type: DNA
 A:Residues: 1-206 <WAI>
 A:Cross-references: GB:K02013; NID:G326417; PIDN:AAB59752.1; PID:G326425
 A:Experimental source: isolate LAV-1a
 R:Ciccarilli, R. B.
 submitted to the EMBL Data Library, March 1991
 A:Reference number: S14607
 A:Accession: S14609
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-206 <CIC>
 A:Cross-references: EMBL:X58780; NID:G60113; PIDN:CAA41585.1; PID:G60114
 C:Genetics:
 A:Gene: nef; 3'-orf; orf-F
 C:Superfamily: AIDS nef protein
 C:Keywords: AIDS; immunodeficiency; phosphoprotein

Query Match 46.1%; Score 83; DB 1; Length 206;
 Best Local Similarity 100.0%; Pred. No. 0.00025;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMTYK 34
 |||||
 Db 68 PPVTPQVPLRPMTYK 82

RESULT 9
 S03244
 nef protein (clone HXB2) - human immunodeficiency virus type 1
 N:Alternate names: 3'-orf protein
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 30-Sep-1993
 C:Accession: S03244
 R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, E.P.; Livak, K.J.; Petteway
 Nucleic Acids Res. 13, 8219-8229, 1985
 A:Title: Polymorphism of the 3' open reading frame of the virus associated with the acqu
 A:Reference number: S03244; MUID:86067228; PMID:2999715
 A:Accession: S03244
 A:Molecule type: DNA
 A:Residues: 1-206 <RAT>
 A:Cross-references: EMBL:X03187
 A:Note: the in-frame stop codon at residue 124, shown as 'X', may be suppressed
 C:Genetics:
 A:Gene: nef; 3'-orf; orf-F
 C:Superfamily: AIDS nef protein

Query Match 46.1%; Score 83; DB 2; Length 206;
 Best Local Similarity 100.0%; Pred. No. 0.00025;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMTYK 34
 |||||
 Db 68 PPVTPQVPLRPMTYK 82

RESULT 10
 S33986
 nef protein - human immunodeficiency virus type 1
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Sep-1999
 C:Accession: S33986
 R:Carlini, F.
 submitted to the EMBL Data Library, November 1991
 A:Reference number: S33979

A:Accession: S33986
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-206 <CAR>
 A:Cross-references: EMBL:Z11530; NID:G60192; PIDN:CAA7629.1; PID:G60200
 C:Superfamily: AIDS nef protein

Query Match 46.1%; Score 83; DB 2; Length 206;
 Best Local Similarity 100.0%; Pred. No. 0.00025;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMTYK 34
 |||||
 Db 68 PPVTPQVPLRPMTYK 82

RESULT 11
 JC5400
 nef protein - Human immunodeficiency virus type 1, HIV-1
 C:Species: Human immunodeficiency virus type 1, HIV-1
 C:Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 31-Oct-1997
 C:Accession: JC5400
 R:Macrae, I.G.; Lowe, M.G.; Curtain, C.C.; Hewish, D.; Azad, A.A.
 Biochem. Biophys. Res. Commun. 232, 707-711, 1997
 A:Title: Cytotoxicity resulting from addition of HIV-1 Nef N-terminal peptides to yeast
 A:Reference number: JC5400; MUID:97271389; PMID:9126340
 A:Accession: JC5400
 A:Molecule type: protein
 A:Residues: 1-206 <MAC>
 C:Comment: The amino-terminal part possesses membrane-perturbing and fusogenic activitie
 F:2-206/Product: nef protein #status predicted <MAT>

Query Match 46.1%; Score 83; DB 2; Length 206;
 Best Local Similarity 100.0%; Pred. No. 0.00025;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMTYK 34
 |||||
 Db 68 PPVTPQVPLRPMTYK 82

RESULT 12
 S25937
 nef protein - human immunodeficiency virus type 1
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 25-Feb-1994 #sequence_revision 01-Dec-1995 #text_change 20-Sep-1999
 C:Accession: S25937
 R:Guo, H.G.; Chernann, J.C.; Waters, D.; Hall, L.; Louie, A.; Gallo, R.C.; Streicher, H.
 Nature 349, 745-746, 1991
 A:Title: Sequence analysis of original HIV-1.
 A:Reference number: S25937; MUID:91156044; PMID:2000145
 A:Accession: S25937
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: nucleic acid
 A:Residues: 1-206 <GUO>
 A:Cross-references: EMBL:X57465; NID:G60217; PIDN:CAA40702.1; PID:G60218
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991
 C:Superfamily: AIDS nef protein

Query Match 43.9%; Score 79; DB 2; Length 206;
 Best Local Similarity 78.9%; Pred. No. 0.00094;
 Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMTYK 34
 |||||
 Db 64 EKVGFPVTPQVPLRPMTYK 82

RESULT 13
 S54385
 nef protein - human immunodeficiency virus type 1 (fragment)

C;Species: human immunodeficiency virus type 1, HIV-1
 C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 20-Sep-1999
 C;Accession: S54385
 R;Theodore, T.; Buckler-White, A.J.
 submitted to the EMBL Data Library, July 1989
 A;Reference number: S54377
 A;Accession: S54385
 A;Status: preliminary
 A;Molecule type: genomic RNA
 A;Residues: 1-97 <THE>
 A;Cross-references: EMBL:M22639; NID:G329377; PIDN:AAA45371.1; PID:G555045
 C;Superfamily: AIDS nef protein

Query Match 43.3%; Score 78; DB 2; Length 97;
 Best Local Similarity 75.0%; Pred. No. 0.00056;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFPVTPQVPLRPMTYK 34
 :| |||||
 Db 64 SEVGFVPRPQVPLRPMTYK 83

RESULT 14

QOLJND
 nef protein - human immunodeficiency virus type 1 (isolate NDK)

N;Alternate names: 3'-orf protein; orf-F protein
 C;Species: human immunodeficiency virus type 1, HIV-1

A;Note: host Homo sapiens (man)

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999

C;Accession: J00068

R;Spite, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe, A.;
 Gene 81, 275-284, 1989

A;Title: Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the human immuno

A;Reference number: J00065; WUID:90034200; PMID:2806917

A;Accession: J00068

A;Molecule type: DNA

A;Residues: 1-207 <SPI>

A;Cross-references: GB:M27323; NID:G328154; PIDN:AAA44874.1; PID:G328163

C;Genetics:

A;Gene: nef

C;Superfamily: AIDS nef protein

C;Keywords: AIDS; immunodeficiency

Query Match 43.3%; Score 78; DB 1; Length 207;
 Best Local Similarity 75.0%; Pred. No. 0.0013;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFPVTPQVPLRPMTYK 34
 :| |||||
 Db 64 SEVGFVPRPQVPLRPMTYK 83

RESULT 15

QOLJZR

nef protein - human immunodeficiency virus Zr-6

N;Alternate names: 3'-orf protein; orf-F protein

C;Species: human immunodeficiency virus Zr-6

C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999

C;Accession: F26192

R;Srinivasan, A.; Anand, R.; Ranganathan, P.; Feorino, P.; Schochetman, G.; Cu

Gene 52, 71-82, 1987

A;Title: Molecular characterization of human immunodeficiency virus from Zaïre: nucleoti

A;Reference number: A26192; WUID:87248097; PMID:3036660

A;Accession: F26192

A;Molecule type: DNA

A;Residues: 1-212 <SRI>

A;Cross-references: GB:K03458; GB:M16322; NID:G329398; PIDN:AAA45381.1; PID:G329404

C;Genetics:

A;Gene: nef; 3'-orf; orf-F

C;Superfamily: AIDS nef protein

C;Keywords: AIDS; immunodeficiency

Query Match 43.3%; Score 78; DB 1; Length 212;

Best Local Similarity 75.0%; Pred. No. 0.0013;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFPVTPQVPLRPMTYK 34
 :| |||||
 Db 69 SEVGFVPRPQVPLRPMTYK 88

Search completed: July 20, 2004, 06:34:10
 Job time : 11.7049 secs


```

Db      64 EKVGFPVTPQVPLRPMPTYK 82

RESULT 2
ID_NEF_HV1PV STANDARD; PRT; 206 AA.
AC P03405;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentiviruses.
OX NCBI_TaxID=11700;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511157; PubMed=2982104;
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Laeky L.A.,
RA Capon D.J.;
RT "Nucleic acid structure and expression of the human
RL AIDS/Lymphadenopathy retrovirus.";
RN Nature 313:450-458(1985).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., Le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RL an oncogene product.";
RL Nature 330:266-269(1987).
CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL; K02083; AAB59874.1; -.
DR EMBL; X01762; -. NOT_ANNOTATED_CDS.
DR PIR; A04007; ASLVL.
DR HSSP; P03406; IEFN.
DR HIV; K02083; NEFSPV22.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
KW AIDS; Myristate; GTP-binding; Lipoprotein.
FT LIPID 2 N-myristoyl glycine (in host).
SQ SEQUENCE 206 AA; 23352 MW; ED664281A17C6735 CRC64;

Query Match 47.2%; Score 85; DB 1; Length 206;
Best Local Similarity 84.2%; Pred. No. 3,1e-05;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db      16 ERGRFPVTPQVPLRPMPTYK 34
ID_NEF_HV1BI STANDARD; PRT; 123 AA.
AC P03404;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).

```

```

CC CCX Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxId=11678;
RN RN [1]
RX RX MEDLINE=6511123; PubMed=2578615.
RA RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Joseph S.F., Doran E.R., Ratacki J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Laemmleberger J.C., Papas T.S., Chirayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL RL Nature 313:277-284(1985).
RP RP [2]
RX RX POST-TRANSLATIONAL MODIFICATIONS AND FUNCTION.
MEDLINE=88039140; PubMed=1118220;
RA Guy B., Kiemy M.-P., Riviere Y., le Peuch C., Doot K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT RT "HIV F/3' ori encodes a phosphorylated GTP-binding protein resembling
an oncogene product.";
RL RL Nature 330:266-269(1987).
CC CC -I- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
activities. It seems to down-regulate the CD4(T4) antigen.
CC CC -I- MISCELLANEOUS: IN THIS ISOLATE A MUTATION IN POSITION 124 ADDS A
STOP CODON AND THE PROTEIN IS MUCH SHORTER THAN IN OTHER ISOLATES
(210 AA).
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC CC or send an email to license@db-sib.ch).
-----
DR DR EMBL; M15654; AAA44206.1; -.
DR DR PIR; A04005; ASLJH3.
DR DR HIV; M15654; NEFEBH102.
DR DR InterPro: IPR001558; HIV_Nef.
DR DR Pfam; PF00469; F-protein; 1.
DR DR ProDom; PD000031; HIV_Nef; 1.
DR DR AIDS; Myristate; GTP-binding; Lipoprotein.
KW KW Lipid 2 N-myristoyl glycine (in host).
FT FT Lipid 2
SQ SEQUENCE 123 AA; 13606 MW; 0811735345FOE8B8 CRC64;

Query Match 46.1%; Score 83; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. NO. 3.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMITYK 34
Db 68 PPVTPQVPLRPMITYK 82

RESULT 4
NEF_HVIH2
ID ID NEF_HVIH2 STANDARD; PRT; 123 AA.
AC AC P04601; O09780;
DT DT 13-AUG-1987 (Rel. 05, Created)
DT DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DE Negative factor (F-protein) (27 Kda protein) (3'ORF).
GN GN NEF.
OS OS Human immunodeficiency virus type 1 (HXB2 isolate) (HTV-1).
OC OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX OX NCBI_TaxId=11706;
RN RN (1)
RP RP SEQUENCE FROM N.A.
RX RX MEDLINE=87929196; PubMed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.T., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT RT *Complete nucleotide sequences of functional clones of the AIDS
RT RT virus.*
RL RL AIDS Res. Hum. Retroviruses 3:57-69(1987).

```

```

RN [2]
RP REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Lion R.-S.,
RA Gallo R.C., Wong-Staal F.,
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.,
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product."
RL Nature 330:266-269(1987).
CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -!- MISCELLANEOUS: IN THIS ISOLATE A MUTATION IN POSITION 124 ADDS A
CC STOP CODON AND THE PROTEIN IS MUCH SHORTER THAN IN OTHER ISOLATES
CC (210 AA).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03455; AAB50263.1; -
DR EMBL; AF03819; AAC82597.1; -.
DR HIT; K03455; NEFSHB2.
DR InterPro; IPR001558; HIV Nef.
DR Pfam; PF00469; F-Protein; 1.
DR ProDom; PD000031; HIV Nef; 1.
KW AIDS; Myristate; GTP-Binding; Phosphorylation; Lipoprotein.
FT LIPID 2 2
FT MOD_RES 15 15 N-myristoyl glycine (in host).
FT PHOSPHORYLATION (BY PKC).
SQ SEQUENCE 123 AA; 13692 MW; B5007753CCD244CF CRC64;

Query Match 46.1%; Score 83; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMTYK 34
DB 68 PPVTPQVPLRPMTYK 82
|||||
|

RESULT 5
ID NEF_HV1B8 STANDARD; PRT; 205 AA.
AC P05855;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1), and
OS Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11684;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Kafaiski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R., Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Ghayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284(1985).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,

```

```

RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product."
RL Nature 330:266-269(1987).
CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02011; -; NOT_ANNOTATED_CDS.
DR HSSP; P03406; 1ERN.
DR HIT; K02011; NEFSB8.
DR InterPro; IPR001558; HIV Nef.
DR Pfam; PF00469; F-Protein; 1.
DR ProDom; PD000031; HIV Nef; 1.
KW AIDS; Myristate; GTP-Binding; Lipoprotein.
FT LIPID 2 2
FT MOD_RES 205 AA; 23305 MW; 8EC12F6650DD111 CRC64;
SQ SEQUENCE 205 AA; 23305 MW; 8EC12F6650DD111 CRC64;

Query Match 46.1%; Score 83; DB 1; Length 205;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMTYK 34
DB 68 PPVTPQVPLRPMTYK 82
|||||
|

RESULT 6
ID NEF_HV1B8 STANDARD; PRT; 206 AA.
AC P03406;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1), and
OS Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11686; 11698;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Isolate BRU;
RX MEDLINE=85099333; PubMed=2981635;
RA Wain-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.,
RT "Nucleotide sequence of the AIDS virus, LAV."
RL Cell 40:9-17(1985).
RN [2]
RP SEQUENCE FROM N.A. (CLONE PNL4-3).
RC STRAIN=Isolate New York-5;
RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;
RL Submitted (JUN-1988) to the EMBL/GenBank/DBJ databases.
RN [3]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.,
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product."
RL Nature 330:266-269(1987).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 54-205 IN COMPLEX WITH NEF.
RX MEDLINE=96279837; PubMed=8681387;
RA Ise C.H., Saksela K., Witza U.A., Chait B.T., Kuriyan J.;
RT "Crystal structure of the conserved core of HIV-1 Nef complexed with
RT a Src family SH3 domain."

```

```

RL Cell 85:931-942(1996).
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02013; AAB59752.1; -
CC EMBL; M19921; AAA44993.1; -
CC EMBL; A04321; CAA00353.1; -
CC PIR; A04008; ASLJFV.
CC
CC PDB; 1EFN; 11-JAN-97.
CC PDB; 1AVZ; 25-MAR-98.
CC PDB; 1AVZ; 25-MAR-98.
CC PDB; 1QA4; 26-MAY-99.
CC DR HIV; K02013; NEFSHRU.
CC DR HIV; M19921; NEFSML43.
CC DR InterPro; IPR001558; HIV_Nef.
CC DR Pfam; PF00469; F-protein; 1.
CC DR ProDom; PD000031; HIV_Nef; 1.
CC DR AIDS; Myristate; GTP-binding; Phosphorylation; 3D-structure;
CC Lipoprotein.
CC
CC KW Lipid
CC FT MOD RES 2 2 N-myristoyl glycine (in host).
CC FT VARIANT 11 15 PHOSPHORYLATION (BY PKC).
CC FT VARIANT 15 15 V -> I (IN CLONE PML4-3).
CC FT VARIANT 33 33 T -> A (IN CLONE PML4-3).
CC FT VARIANT 51 51 A -> V (IN CLONE PML4-3).
CC FT VARIANT 51 51 T -> N (IN CLONE PML4-3).
CC
CC FT HELIX 81 93
CC FT TURN 94 94
CC FT STRAND 98 98
CC FT TURN 99 99
CC FT STRAND 101 101
CC FT HELIX 104 118
CC FT STRAND 127 127
CC FT STRAND 134 134
CC FT STRAND 136 136
CC FT TURN 139 140
CC FT STRAND 143 147
CC FT STRAND 181 185
CC FT HELIX 187 190
CC FT TURN 191 191
CC FT HELIX 194 198
CC FT HELIX 200 202
CC
CC SQ SEQUENCE 206 AA; 23342 MW; 77453FC80B6004F2 CRC64;
CC
CC Query Match 46.1%; Score 83; DB 1; Length 206;
CC Best Local Similarity 100.0%; Pred. No. 6.2e-05;
CC Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 20 FPVTPQVPLRPMYK 34
CC DB 68 FPVTPQVPLRPMYK 82
CC
CC RESULT 7
CC NEF_HV11W STANDARD; PRT; 206 AA.
CC ID NEF_HV11W
CC AC Q70627;
CC DT 15-JUL-1998 (Rel. 36, Created)
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DB Negative Factor (F-protein) (27 kDa protein) (3'ORF).
CC GN NEF.
CC OS Human immunodeficiency virus type 1 (HIV-1).
CC OC Viruses; Retroviridae; Lentivirinae.
CC OX NCBI_TaxID=82834;
CC RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M.S. et al., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,
RA Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker
RT infected with HIV type 1 (HTLV type IIb)".
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP STRUCTURE BY NMR OF 56-206.
RX MEDLINE=97337445; PubMed=9194185;
RA Grzesiek S., Bax A., Hu U.S., Kaufman J., Palmer I., Stahl S.J.,
RA Tjandra N., Wingfield P.T.;
RT "Refined solution structure and backbone dynamics of HIV-1 Nef.";
RL Protein Sci. 6:1248-1263(1997).
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U12055; AAA76691.1; -
CC PDB; 2NEF; 07-JUL-97.
CC DR InterPro; IPR001558; HIV_Nef.
CC DR Pfam; PF00469; F-protein; 1.
CC DR ProDom; PD000031; HIV_Nef; 1.
CC KW AIDS; Myristate; GTP-binding; 3D-structure; Lipoprotein.
CC FT LIPID 2 2 N-myristoyl glycine (in host).
CC SQ SEQUENCE 206 AA; 23414 MW; ED81F68F6B61278E CRC64;
CC
CC Query Match 46.1%; Score 83; DB 1; Length 206;
CC Best Local Similarity 100.0%; Pred. No. 6.2e-05;
CC Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 20 FPVTPQVPLRPMYK 34
CC DB 68 FPVTPQVPLRPMYK 82
CC
CC RESULT 8
CC NEF_HV122 STANDARD; PRT; 97 AA.
CC ID NEF_HV122
CC AC P12478;
CC DT 01-OCT-1989 (Rel. 12, Created)
CC DT 01-OCT-1989 (Rel. 12, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DB Negative Factor (F-protein) (27 kDa protein) (3'ORF) (Fragment).
CC GN NEF.
CC OS Human immunodeficiency virus type 1 (22/CDC-234 isolate) (HIV-1).
CC OC Viruses; Retroviridae; Lentivirinae.
CC OX NCBI_TaxID=11683;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RA Theodore T., Buckler-White A.;
CC Submitted (NOV-1988) to the HIV data bank.
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M22639; AAA45371.1; -
CC PIR; S54385; S54385.
CC DR HIV; M22639; NEF$2266.

```



```

DR InterPro: IPR001558; HIV_Nef.
DR Pfam: PF00469; F-protein. 1.
DR Prodom: PD000031; HIV_Nef. 1.
KM AIDS; Myristate; GTP-Binding; Lipoprotein.
FT LIPID 2 N-myristoyl glycine (in host) (By
FT NON TER 97 97 similarity).
SQ SEQUENCE 97 AA; 10592 MW; F8529A065B670AE9 CRC64;

Query Match 43.3%; Score 78; DB 1; Length 97;
Best Local Similarity 75.0%; Pred. No. 0.00015;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFVPTQVPLRPMTYK 34
Db 64 SEVGFVPRQVPLRPMTYK 83

RESULT 9
ID NEF_HV1ND STANDARD; PRT; 207 AA.
AC P18601;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OK NCBI_TaxID=11695;
RX MEDLINE=90034200; PubMed=2806917;
RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,
RA Hampe A., Chermann J.C.;
RT Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the
RT human immunodeficiency virus."
RL Gene 81:275-284(1989).
CC CC
CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(74) antigen.
CC -!- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M27323; AAA44874.1; -.
DR PIR; J00068; Q0LND.
DR HSSP; P03406; IEFN.
DR HIV; M27323; NEFNDK.
DR InterPro: IPR001558; HIV_Nef.
DR Pfam: PF00469; F-protein. 1.
DR Prodom: PD000031; HIV_Nef. 1.
KM AIDS; Myristate; GTP-Binding; Lipoprotein.
FT LIPID 2 N-myristoyl glycine (in host) (By
FT NON TER 97 97 similarity).
SQ SEQUENCE 207 AA; 23748 MW; 09036C2F81D45D5E CRC64;

Query Match 43.3%; Score 78; DB 1; Length 207;
Best Local Similarity 75.0%; Pred. No. 0.00034;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFVPTQVPLRPMTYK 34
Db 64 SEVGFVPRQVPLRPMTYK 83

RESULT 10

```

```

NEF_HV1Z6
ID NEF_HV1Z6 STANDARD; PRT; 212 AA.
AC P04602;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OK NCBI_TaxID=11708;
RX MEDLINE=87248097; PubMed=3036660;
RA Sriivasan A., Anand R., York D., Ranganathan P., Feorino P.,
RA Schocherman G., Curran J., Kalyanaraman V.S., Luciw P.A.,
RA Sanchez-Pescador R.;
RT "Molecular characterization of human immunodeficiency virus from
RT Zaire: nucleotide sequence analysis identifies conserved and variable
RT domains in the envelope gene".
RL Gene 52:71-82(1987).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dolt K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product".
RL Nature 330:266-269(1987).
CC CC
CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(74) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; K03458; AAA45381.1; -.
DR PIR; F26192; Q0LZR.
DR HSSP; P03406; IEFN.
DR HIV; K03458; NEFZ6.
DR InterPro: IPR001558; HIV_Nef.
DR Pfam: PF00469; F-protein. 1.
DR Prodom: PD000031; HIV_Nef. 1.
KM AIDS; Myristate; GTP-Binding; Lipoprotein.
FT LIPID 2 N-myristoyl glycine (in host).
FT NON TER 212 212 EC095A1B263047CF CRC64;
SQ SEQUENCE 212 AA; 24415 MW; EC095A1B263047CF CRC64;

Query Match 43.3%; Score 78; DB 1; Length 212;
Best Local Similarity 75.0%; Pred. No. 0.00035;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFVPTQVPLRPMTYK 34
Db 69 SEVGFVPRQVPLRPMTYK 88

RESULT 11
ID NEF_HV1MN STANDARD; PRT; 182 AA.
AC P05856;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OK NCBI_TaxID=11696;
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dolt K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product".
RL Nature 330:266-269(1987).
CC CC
CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(74) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; K03458; AAA45381.1; -.
DR PIR; F26192; Q0LZR.
DR HSSP; P03406; IEFN.
DR HIV; K03458; NEFZ6.
DR InterPro: IPR001558; HIV_Nef.
DR Pfam: PF00469; F-protein. 1.
DR Prodom: PD000031; HIV_Nef. 1.
KM AIDS; Myristate; GTP-Binding; Lipoprotein.
FT LIPID 2 N-myristoyl glycine (in host).
FT NON TER 212 212 EC095A1B263047CF CRC64;
SQ SEQUENCE 212 AA; 24415 MW; EC095A1B263047CF CRC64;

Query Match 43.3%; Score 78; DB 1; Length 212;
Best Local Similarity 75.0%; Pred. No. 0.00035;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFVPTQVPLRPMTYK 34
Db 69 SEVGFVPRQVPLRPMTYK 88

RESULT 11

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=86219542; PubMed=3369091;
RA Guigo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates.";
FT Virology 164:531-536(1988).
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -1- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
CC PATIENT IN 1984.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17449; AAA44858.1; -.
DR HSSP; P03406; IEFN.
DR HIV; M17449; NEFSMN.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
DR AIDS; Myristate; GTP-binding; Lipoprotein.
DR AIDS; Myristate; GTP-binding; Lipoprotein.
FT Lipid 2
FT N-myristoyl glycine (in host) (By
FT similarity).
SQ SEQUENCE 182 AA; 20632 MW; 2C5B5A6E3EF13B26 CRC64;

Query Match 42.8%; Score 77; DB 1; Length 182;
Best Local Similarity 93.3%; Pred. No. 0.00048;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 FVTPQVPLRPMTYK 34
DB 70 FVTPQVPLRPMTYK 84

RESULT 12
NEF_HV1S3 STANDARD; PRT; 205 AA.
AC P19545;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).
OS Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11690;
RN (1)
RS SEQUENCE FROM N.A.
RX MEDLINE=90317906; PubMed=2370688;
RA York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;
RT "Human immunodeficiency virus type 1 cellular host range,
RT replication, and cytopathicity are linked to the envelope region of
RT the viral genome.";
FT J. Virol. 64:4016-4020(1990).
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AY352275; AAQ17032.1; -.
DR HSSP; P03406; IEFN.
DR HIV; M38427; NEFSF33.

```

```

DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
DR AIDS; Myristate; GTP-binding; Lipoprotein.
FT Lipid 2
FT N-myristoyl glycine (in host) (By
FT similarity).
SQ SEQUENCE 205 AA; 23318 MW; FFB419A1C5DFC9F3 CRC64;

Query Match 42.8%; Score 77; DB 1; Length 205;
Best Local Similarity 93.3%; Pred. No. 0.00048;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 FVTPQVPLRPMTYK 34
DB 67 FVTPQVPLRPMTYK 81

RESULT 13
NEF_HV1U4 STANDARD; PRT; 205 AA.
AC P24741;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (strain Ugandan / isolate U455)
OS (HIV-1).
OS Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11703;
RN (1)
RS SEQUENCE FROM N.A.
RX MEDLINE=91090981; PubMed=2265025;
RA Oram J.D., Downing R.G., Roff M., Clegg J.C.S., Serwadda D.,
RA Carswell J.W.;
RT "Nucleotide sequence of a Ugandan HIV-1 provirus reveals genetic
RT diversity from other HIV-1 isolates.";
RT AIDS Res. Hum. Retroviruses 6:1073-1078(1990).
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62320; AAA75023.1; -.
DR HSSP; P03406; IEFN.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
DR AIDS; Myristate; GTP-binding; Lipoprotein.
FT Lipid 2
FT N-myristoyl glycine (in host) (By
FT similarity).
SQ SEQUENCE 205 AA; 23253 MW; 869AB03B6E7893C4 CRC64;

Query Match 42.8%; Score 77; DB 1; Length 205;
Best Local Similarity 93.3%; Pred. No. 0.00048;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 FVTPQVPLRPMTYK 34
DB 68 FVTPQVPLRPMTYK 82

RESULT 14
NEF_HV1EL STANDARD; PRT; 206 AA.
AC P04604;
DT 13-AUG-1987 (Rel. 05, Created)

```

```

DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (EDI isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
RL of two isolates from African patients.";
RN Cell 46:63-74(1986).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dett K., Girard M.,
RT "HIV F3' orf encodes a phosphorylated GTP-binding protein resembling
an oncogene product.";
RL Nature 330:266-269(1987).
CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
activities. It seems to down-regulate the CD4(T4) antigen.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; K03454; AAA44330.1; -
DR EMBL; A07108; CAA00617.1; -
DR HSSP; P03406; 1EEN.
DR HIV; K03454; NEPSL1.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-Protein_1.
DR ProDom; PD000031; HIV_Nef; 1.
DR AIDS; Myristate; GTP-binding; Lipoprotein.
DR LIPID 2
FT SEQUENCE 206 AA; 23612 MW; 4CFE9F18AEB503C CRC64;

Query Match 42.8%; Score 77; DB 1; Length 206;
Best Local Similarity 93.3%; Pred. No. 0.00048;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMTYK 34
Db 69 PPVTPQVPLRPMTYK 83

RESULT 15
NEF_HV1S1 STANDARD; PRT; 208 AA.
ID NEF_HV1S1
AC P19546;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (SP162 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating

```

```

CC activities. It seems to down-regulate the CD4(T4) antigen.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M65024; AAA5073.1; -
DR HSSP; P03406; 1EEN.
DR HIV; M38428; NEPSF162.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-Protein_1.
DR ProDom; PD000031; HIV_Nef; 1.
DR AIDS; Myristate; GTP-binding; Lipoprotein.
DR LIPID 2
FT SEQUENCE 208 AA; 23684 MW; A0B1007D14E4E32 CRC64;

Query Match 42.8%; Score 77; DB 1; Length 208;
Best Local Similarity 93.3%; Pred. No. 0.00048;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMTYK 34
Db 70 PPVTPQVPLRPMTYK 84

```

Search completed: July 20, 2004, 06:31:02
Job time : 7.68852 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:43 ; Search time 34.2787 Seconds

(without alignments)
312.953 Million cell updates/sec

Title: US-09-673-166B-276
Perfect score: 180

Sequence: 1 GRQYIKANSKFIGITERGRFPVTPQVPLRPMYTK 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.5	48.1	142	15	Q8JB99 human immun
2	86	47.8	141	15	Q8JE98 human immun
3	86	47.8	142	15	Q8JE95 human immun
4	85.5	47.5	208	15	Q74912 human immun
5	85	47.2	142	15	Q8JB97 human immun
6	85	47.2	142	15	Q8JB96 human immun
7	85	47.2	206	15	Q8JS88 human immun
8	84	46.7	208	15	Q8JS88 aids-associ
9	83.5	46.4	207	15	P90287 human immun
10	83	46.1	90	15	P90287 human immun
11	83	46.1	118	15	Q41584 human immun
12	83	46.1	121	15	Q79786 human immun
13	83	46.1	123	15	Q85587 aids-associ
14	83	46.1	153	15	Q74916 human immun
15	83	46.1	176	15	Q74918 human immun
16	83	46.1	202	15	Q8G617 human immun

17	83	46.1	202	15	Q9GPN3 human immun
18	83	46.1	204	15	Q9GSC4 human immun
19	83	46.1	204	15	Q9WIP2 human immun
20	83	46.1	204	15	Q9WIP3 human immun
21	83	46.1	204	15	Q74920 human immun
22	83	46.1	204	15	Q74921 human immun
23	83	46.1	204	15	Q74921 human immun
24	83	46.1	204	15	Q76352 human immun
25	83	46.1	204	15	Q75628 human immun
26	83	46.1	204	15	Q72FD3 human immun
27	83	46.1	205	15	Q9TH66 human immun
28	83	46.1	205	15	Q9DQT4 human immun
29	83	46.1	206	15	Q9DQT8 human immun
30	83	46.1	206	15	Q9DQT8 human immun
31	83	46.1	206	15	Q9MM30 human immun
32	83	46.1	206	15	Q8J512 human immun
33	83	46.1	206	15	Q8J512 human immun
34	83	46.1	206	15	Q9DQT9 human immun
35	83	46.1	206	15	Q9DQT9 human immun
36	83	46.1	206	15	Q9DQT9 human immun
37	83	46.1	206	15	Q9DQT9 human immun
38	83	46.1	206	15	Q9DQT9 human immun
39	83	46.1	206	15	Q9DQT9 human immun
40	83	46.1	206	15	Q9DQT9 human immun
41	83	46.1	206	15	Q9DQT9 human immun
42	83	46.1	206	15	Q9DQT9 human immun
43	83	46.1	206	15	Q9DQT9 human immun
44	83	46.1	206	15	Q9DQT9 human immun
45	83	46.1	206	15	Q9DQT9 human immun

ALIGNMENTS

RESULT 1

Q8JB99 PRELIMINARY; PRT; 142 AA.
ID Q8JB99
AC Q8JB99;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Envelope glycoprotein (Negative factor) (F-protein) (27 kDa protein) (Fragment).
DE ENV.
GN Human immunodeficiency virus 1.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeuchi H., Suzuki Y., Hoshino H., Daar E.S., Koyanagi Y.,
RT "Extensive variation of CXCR4- and CCR5-dependent human
RT immunodeficiency virus type 1 in a patient with acute infection."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
CC EMBL; AB086306; BAC02668.1; -.
DR GO; GO:0005525; F-GTP binding; IRA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
KW AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON TER 142
SQ SEQUENCE 142 AA; 15727 MW; 152F51521B88A13 CRC64;

Query Match 48.1%; Score 86.5; DB 15; Length 142;
Best Local Similarity 58.1%; Pred. No. 0.0001;
Matches 18; Conservative 4; Mismatches 4; Indels 5; Gaps 1;

QY 4 YIKANSKFIGITERGRFPVTPQVPLRPMYTK 34
DB 57 WIKANQER-----EKVGFPVTPQVPLRPMYTK 82

RESULT 2

```
08J98      PRELIMINARY;      PRT;      141 AA.
AC 08J98;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Envelope glycoprotein (Negative factor) (F-protein) (27 kDa protein)
DE (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeuchi H., Suzuki Y., Hoshino H., Daar E.S., Koyanagi Y.;
RT "Extensive variation of CXCR4- and CCR5-dependent human
immunodeficiency virus type 1 in a patient with acute infection.";
Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN (BY
SIMILARITY).
DR EMBL; AB086307; BAC02669.1; -.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-Protein; 1.
DR Prodom; PD000031; HIV_Nef; 1.
DR AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON_TER
SQ SEQUENCE 141 AA; 15544 MW; 2DD72C4B9C27097 CRC64;

Query Match      47.8%; Score 86; DB 15; Length 141;
Best Local Similarity 84.2%; Pred. No. 0.00012;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      16 ERGRFPVTPQVPLRPMYTK 34
Db      64 EKVSFPVTPQVPLRPMYTK 82

RESULT 3
08J95      PRELIMINARY;      PRT;      142 AA.
AC 08J95;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Envelope glycoprotein (Negative factor) (F-protein) (27 kDa protein)
DE (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeuchi H., Suzuki Y., Hoshino H., Daar E.S., Koyanagi Y.;
RT "Extensive variation of CXCR4- and CCR5-dependent human
immunodeficiency virus type 1 in a patient with acute infection.";
Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN (BY
SIMILARITY).
DR EMBL; AB086310; BAC02672.1; -.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR Prodom; PD000031; HIV_Nef; 1.
DR AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON_TER
SQ SEQUENCE 142 AA; 15827 MW; 09C2333F280A0FDF CRC64;

Query Match      47.8%; Score 86; DB 15; Length 142;
```

Best Local Similarity 84.2%; Pred. No. 0.00012;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
QY      16 ERGRFPVTPQVPLRPMYTK 34
Db      64 EKVSFPVTPQVPLRPMYTK 82
```

RESULT 4

```
074912     PRELIMINARY;      PRT;      208 AA.
AC 074912;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Nef (Negative factor) (F-protein) (27 kDa protein).
GN NEF.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96400183; PubMed=8806559;
RA Rather L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,
RA Hahn B., Powderly W., Arens M.;
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects
at different stages of disease.";
Virology 223:245-250(1996).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN (BY
SIMILARITY).
DR EMBL; U04452; AAB38204.1; -.
DR HSP; P03406; IEA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-Protein; 1.
DR Prodom; PD000031; HIV_Nef; 1.
DR AIDS; GTP-binding; Lipoprotein; Myristate.
SQ SEQUENCE 208 AA; 23596 MW; 04B09B3B594DBA0 CRC64;
```

Query Match 47.5%; Score 85.5; DB 15; Length 208;
Best Local Similarity 58.1%; Pred. No. 0.00022;
Matches 18; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

```
QY      7 ANSKFVGITERGR--FPVTPQVPLRPMYTK 34
Db      54 ADCAWVEAQEGEGVGFVTPQVPLRPMYTK 84
```

RESULT 5

```
08J97      PRELIMINARY;      PRT;      142 AA.
AC 08J97;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Envelope glycoprotein (Negative factor) (F-protein) (27 kDa protein)
DE (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeuchi H., Suzuki Y., Hoshino H., Daar E.S., Koyanagi Y.;
RT "Extensive variation of CXCR4- and CCR5-dependent human
immunodeficiency virus type 1 in a patient with acute infection.";
Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN (BY
SIMILARITY).
DR EMBL; AB086308; BAC02670.1; -.
DR GO; GO:0005525; F:GTP binding; IEA.
```

```
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
KM AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON_TER 142
SQ SEQUENCE 142 AA; 15715 MW; 9BC3275290AEFB7D CRC64;

Query Match 47.2%; Score 85; DB 15; Length 142;
Best Local Similarity 84.2%; Pred. No. 0.00017;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMYTK 34
   |:|||||
Db 64 EKVGFPVTPQVPLRPMYTK 82

RESULT 6
ID Q8UE96 PRELIMINARY; PRT; 142 AA.
AC Q8UE96;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein (Negative factor) (F-protein) (27 kDa protein) (Fragment).
DE (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeuchi H., Suzuki Y., Hoshino H., Daar E.S., Koyanagi Y.;
RT "Extensive variation of CXCR4- and CCR5-dependent human
RT immunodeficiency virus type 1 in a patient with acute infection.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; AB086309; BAC02671.1; -
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
KM AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON_TER 142
SQ SEQUENCE 142 AA; 15715 MW; 9BC3275290AEFB7D CRC64;

Query Match 47.2%; Score 85; DB 15; Length 142;
Best Local Similarity 84.2%; Pred. No. 0.00017;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMYTK 34
   |:|||||
Db 64 EKVGFPVTPQVPLRPMYTK 82

RESULT 7
ID Q85588 PRELIMINARY; PRT; 206 AA.
AC Q85588;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T-lymphotropic virus type III (HTLV III) 3' ORF HXB3 RNA (Negative
DE factor) (F-protein) (27 kDa protein).
GN NEF.
OS Aids-associated retrovirus.
OC Viruses; Retroid viruses; Retroviridae.
OX NCBI_TaxID=11966;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86067228; PubMed=2999715;
RA Ratner L., Starcich B., Josephs S.F., Hahn B.H., Reddy E.P.,
```

```
RA Livak K.J., Petteway S.R.Jr., Pearson M.L., Haseltine W.A., Arya S.K.,
RA Wong-staal F.;
RT "Polymorphism of the 3' open reading frame of the virus associated
RT with the acquired immune deficiency syndrome, human T-lymphotropic
RT virus type III ";
RL Nucleic Acids Res. 13:8219-8229(1985).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; X03188; CAA26947.1; -
DR HSSP; P03406; IEFN.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
KM AIDS; GTP-binding; Lipoprotein; Myristate.
SQ SEQUENCE 206 AA; 23419 MW; 3B25EB332A479A46 CRC64;

Query Match 47.2%; Score 85; DB 15; Length 206;
Best Local Similarity 84.2%; Pred. No. 0.00026;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMYTK 34
   |:|||||
Db 64 EKVGFPVTPQVPLRPMYTK 82

RESULT 8
ID O91064 PRELIMINARY; PRT; 208 AA.
AC O91064;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nef protein (Negative factor) (27 kDa protein) (Fragment).
GN NEF.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Saragosti S., Peeters M.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; AJ233029; CAA13510.1; -
DR HSSP; P03406; IEFN.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
KM AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON_TER 208
SQ SEQUENCE 208 AA; 23839 MW; F96EBA1A82BE03C2 CRC64;

Query Match 46.7%; Score 84; DB 15; Length 208;
Best Local Similarity 84.2%; Pred. No. 0.00037;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMYTK 34
   |:|||||
Db 66 EAVRFPVTPQVPLRPMYTK 84

RESULT 9
ID P90287 PRELIMINARY; PRT; 207 AA.
AC P90287;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nef protein (Negative factor) (27 kDa protein).
```

```

GN NEF.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZAM184;
RX MEDLINE=96036482; PubMed=7483282;
RA Salminen M.O., Carr J.K., Sanders-Buell E., Ehrenberg P.K.,
RT Michael N.L., Carr J.K., Burke D.S., McCutchan F.E.;
RT "Recovery of virtually full-length HIV-1 provirus of diverse subtypes
from primary virus cultures using the polymerase chain reaction.";
RN J. Virol. 71:2647-2655(1997).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=ZAM184;
RX MEDLINE=97213928; PubMed=9060617;
RA Salminen M.O., Carr J.K., Robertson D.L., Hegerich P., Gotte D.,
RT Koch C., Sanders-Buell E., Gao F., Sharp P.M., Hahn B.H., Burke D.S.,
RT McCutchan F.E.;
RT "Evolution and probable transmission of intersubtype recombinant human
immunodeficiency virus type 1 in a Zambian couple.";
RN J. Virol. 71:2647-2655(1997).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=ZAM184;
RX Salminen M.O., Carr J.K., Robertson D.L., Hegerich P., Gotte D.,
RA Koch C., Sanders-Buell E., Gao F., Sharp P.M., Hahn B., Burke D.S.,
RT McCutchan F.E.;
RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL: U86780; AAC57010.1; -.
DR HSSP: P03406; 1EFN.
DR GO: GO:0005525; F:GTP binding; IEA.
DR InterPro: IPR001558; HIV_Nef.
DR Pfam: PF00469; F-protein; 1.
DR ProDom: PD000031; HIV_Nef; 1.
DR AIDS: GTP-binding; Lipoprotein; Myristate.
SQ SEQUENCE 207 AA; 23358 MW; A267097BEF3A69F7 CRC64;

Query Match 46.4%; Score 83.5; DB 15; Length 207;
Best Local Similarity 63.0%; Pred. No. 0.00044;
Matches 17; Conservative 2; Mismatches 5; Indels 3; Gaps 1;

QY 11 FVTPQVPLRPMTYK 34
Db 57 WLGAGBGEVGFVKQVFLRPMTYK 83

RESULT 10
ID 041645 PRELIMINARY; PRT; 90 AA.
AC 041645;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Nef protein (Negative factor) (27 kDa protein) (Fragment).
GN NEF.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C18;
RX MEDLINE=98105804; PubMed=9445059;
RA Connor R.I., Korber B.T.M., Graham B.S., Hahn B.H., Ho D.D.,
RA Walker B.D., Neumann A.U., Vermund S.H., Mestecky J., Jackson S.,
RA Fenimore B., Cao Y., Gao F., Kalams S., Kunstman K.J., McDonald D.,
RA McWilliams N., Trkola A., Moore J.P., Wolinsky S.M.;
RT "Immunological and virological analyses of persons infected by human
immunodeficiency virus type 1 while participating in trials of

```

```

RT recombinant gp120 subunit vaccines.";
RL J. Virol. 72:1552-1576(1998).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL: U84886; AAC58933.1; -.
DR GO: GO:0005525; F:GTP binding; IEA.
DR InterPro: IPR001558; HIV_Nef.
DR Pfam: PF00469; F-protein; 1.
DR ProDom: PD000031; HIV_Nef; 1.
DR AIDS: GTP-binding; Lipoprotein; Myristate.
FT NON TER 90
SQ SEQUENCE 90 AA; 9794 MW; A3A713557855DF3A CRC64;

Query Match 46.1%; Score 83; DB 15; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FVTPQVPLRPMTYK 34
Db 68 FVTPQVPLRPMTYK 82

RESULT 11
ID 041584 PRELIMINARY; PRT; 118 AA.
AC 041584;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Nef protein (Negative factor) (27 kDa protein) (Fragment).
GN NEF.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C13;
RX MEDLINE=98105804; PubMed=9445059;
RA Connor R.I., Korber B.T.M., Graham B.S., Hahn B.H., Ho D.D.,
RA Walker B.D., Neumann A.U., Vermund S.H., Mestecky J., Jackson S.,
RA Fenimore B., Cao Y., Gao F., Kalams S., Kunstman K.J., McDonald D.,
RA McWilliams N., Trkola A., Moore J.P., Wolinsky S.M.;
RT "Immunological and virological analyses of persons infected by human
immunodeficiency virus type 1 while participating in trials of
recombinant gp120 subunit vaccines.";
RL J. Virol. 72:1552-1576(1998).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL: U84833; AAC58872.1; -.
DR GO: GO:0005525; F:GTP binding; IEA.
DR InterPro: IPR001558; HIV_Nef.
DR Pfam: PF00469; F-protein; 1.
DR ProDom: PD000031; HIV_Nef; 1.
DR AIDS: GTP-binding; Lipoprotein; Myristate.
FT NON TER 118
SQ SEQUENCE 118 AA; 13186 MW; DD529808140327C2 CRC64;

Query Match 46.1%; Score 83; DB 15; Length 118;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FVTPQVPLRPMTYK 34
Db 70 FVTPQVPLRPMTYK 84

RESULT 12
ID 079786 PRELIMINARY; PRT; 121 AA.
AC 079786;
DT 01-NOV-1996 (TREMBLrel. 01, Created)

```



```

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nef protein (Negative factor) (27 kDa protein).
GN Nef.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCRnef6;
RX MEDLINE=92385154; PubMed=1515209;
RA Harris M., Hislop S., Patelinasos P., Neil J.C.;
RT "In vivo derived HIV-1 Nef genes are heterogeneous and lack detectable
RT nucleotide binding activity."
RL AIDS Res. Hum. Retroviruses 8:537-543(1992).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; X63045; CAA44771.1; -.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein_1.
DR ProDom; PD000031; HIV_Nef; 1.
KW AIDS; GTP-binding; Lipoprotein; Myristate.
SQ SEQUENCE 121 AA; 13506 MW; 0B9226A9052A2CC7 CRC64;

Query Match
Best Local Similarity 46.1%; Score 83; DB 15; Length 121;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMYTK 34
DB 66 FPVTPQVPLRPMYTK 80

RESULT 13
Q85587 PRELIMINARY; PRT; 123 AA.
ID Q85587;
AC Q85587;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T-Lymphocytic virus type III (HTLV-III) 3'ORF HXB2 RNA (Negative
DE factor) (F-protein) (27 kDa protein).
GN Nef.
OS Aids-associated retrovirus.
OC Viruses; Retroid viruses; Retroviridae.
OX NCBI_TaxID=11966;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067228; PubMed=2999715;
RA Ratner L., Starcich B., Josephs S.F., Hahn B.H., Reddy E.P.,
RA Livak K.J., Petteway S.R.Jr., Pearson M.L., Haseltine W.A., Arya S.K.,
RA Wong-staal F.;
RT "Polymorphism of the 3' open reading frame of the virus associated
RT with the acquired immune deficiency syndrome, human T-lymphotropic
RT virus type III."
RL Nucleic Acids Res. 13:8219-8229(1985).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; X03187; CAA26946.1; -.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein_1.
DR ProDom; PD000031; HIV_Nef; 1.
KW AIDS; GTP-binding; Lipoprotein; Myristate.
SQ SEQUENCE 123 AA; 13593 MW; ABB6AC3BE7A2508A CRC64;

Query Match
Best Local Similarity 46.1%; Score 83; DB 15; Length 123;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 20 FPVTPQVPLRPMYTK 34
DB 68 FPVTPQVPLRPMYTK 82

RESULT 14
Q74916 PRELIMINARY; PRT; 153 AA.
ID Q74916;
AC Q74916;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nef (Negative factor) (F-protein) (27 kDa protein)
DE (Fragment).
GN Nef.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96400183; PubMed=8806559;
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,
RA Hahn B., Powderly W., Arens M.;
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects
RT at different stages of disease."
RL Virology 223:245-250(1996).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; U44459; AAB38209.1; -.
DR HSSP; P03406; IEFN.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein_1.
DR ProDom; PD000031; HIV_Nef; 1.
KW AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON_TER 153
SQ SEQUENCE 153 AA; 16939 MW; BAE22A2A21305FDD CRC64;

Query Match
Best Local Similarity 46.1%; Score 83; DB 15; Length 153;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMYTK 34
DB 68 FPVTPQVPLRPMYTK 82

RESULT 15
Q74918 PRELIMINARY; PRT; 176 AA.
ID Q74918;
AC Q74918;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nef (Negative factor) (F-protein) (27 kDa protein)
DE (Fragment).
GN Nef.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96400183; PubMed=8806559;
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,
RA Hahn B., Powderly W., Arens M.;
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects
RT at different stages of disease."
RL Virology 223:245-250(1996).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; U44461; AAB38211.1; -.

```

DR HSSE; P03406; IEFN.
 DR GO; GO:0005525; F:GTP binding; IEA.
 DR InterPro; IPR01558; HIV_Nef.
 DR Pfam; PF00469; F-protein; 1.
 DR ProDom; PD000031; HIV_Nef; 1.
 KW AIDS; GTP-binding; Lipoprotein; Myristate.
 FT NON_TER 176 176
 SQ SEQUENCE 176 AA; 19286 MW; 8EE2DFCBF924430B CRC64;

Query Match 46.1%; Score 83; DB 15; Length 176;
 Best Local Similarity 100.0%; Pred. No. 0.00043;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMTRYK 34
 |||||
 Db 68 FPVTPQVPLRPMTRYK 82

Search completed: July 20, 2004, 06:33:19
 Job time : 35.2787 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:07:17 ; Search time 50.7213 Seconds

(without alignments)
189,400 Million cell updates/sec

Title: US-09-673-166B-276

Perfect score: 180

Sequence: 1 GROYIKANSKFIGITGRGPVTPQVPLRPMYK 34

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched:

1586107 segs, 282547505 residues

Total number of hits satisfying chosen parameters:

1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	53.9	28	2	AA53552 Lipopepti
2	89	49.4	280	2	AAW71642 HIV POL/N
3	89	49.4	280	2	AAW71363 HIV POL/N
4	89	49.4	280	4	AA898912 Synthetic
5	87.5	48.6	25	4	AA884738 Amino aci
6	86	47.8	29	2	AA53551 Lipopepti
7	85	47.2	206	1	AA61515 Sequence
8	85	47.2	206	2	AAW89326 HIV-1 nef
9	85	47.2	206	6	ABU63323 Human lym
10	84	46.7	505	6	ABR44739 Plasmid P
11	84	46.7	505	6	ABR44747 Plasmid T
12	84	46.7	1067	6	ABR44751 Plasmid T
13	84	46.7	1067	6	ABR44753 Plasmid T
14	83	46.1	30	4	AA831115 A polypep
15	83	46.1	30	4	AA830938 Peptide c
16	83	46.1	30	4	AA831009 Polypepti
17	83	46.1	33	2	AA526605 HIV-deriv
18	83	46.1	33	5	ABB83185 Universal
19	83	46.1	37	2	AA653389 Superant
20	83	46.1	38	2	AA74069 HIV Negat
21	83	46.1	38	2	AA739988 HIV Negat
22	83	46.1	46	2	AA426595 Peptide P
23	83	46.1	48	1	AA922268 nef CTL1
24	83	46.1	83	2	AA894760 HIV Nef-1
25	83	46.1	83	4	AA898916 HIV Nef-1

26	83	46.1	83	5	ABB83193 HIV cytot
27	83	46.1	123	4	AA866000 Amino aci
28	83	46.1	142	6	ABR44742 Plasmid 7
29	83	46.1	151	5	AAW52395 Viral fus
30	83	46.1	206	2	AA838893 Nef prote
31	83	46.1	206	2	AAW90179 HTLV-II
32	83	46.1	206	3	AA810054 HIV-1 nef
33	83	46.1	206	3	AA50795 Human NEF
34	83	46.1	206	5	AAW52396 Viral fus
35	83	46.1	206	6	ABU63328 Human lym
36	83	46.1	215	2	AAU02349 A represe
37	83	46.1	215	4	AA63231 Amino aci
38	83	46.1	215	7	ADB67687 HIV-1 nef
39	83	46.1	216	1	AA60423 Sequence
40	83	46.1	216	5	AA019390 Lymphaden
41	83	46.1	239	5	AAW52397 Viral fus
42	83	46.1	302	2	AAU02357 A represe
43	83	46.1	302	2	AAU02351 A represe
44	83	46.1	302	4	AA63239 Amino aci
45	83	46.1	302	4	AA63233 Amino aci

ALIGNMENTS

RESULT 1	AA53552	AA53552 standard; protein; 28 AA.
ID	AA53552	
XX	AA53552;	
AC	18-JAN-2000 (first entry)	
XX		
DT		
XX		
DE	Lipopeptide #3.	
XX		
KW	Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;	
KW	electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;	
KW	human immunodeficiency virus; hepatitis B virus; papilloma virus;	
KW	melanoma; malaria; parasite.	
OS	Synthetic.	
OS	Homo sapiens.	
XX		
FX		
FT	Key	Location/Qualifiers
FT	Modified-site	1
FT	/note= "contains palmitoyl residue attached at the N-	
FT	terminus and on the epsilon carbon of the side chain"	
XX		
PN	FR2776926-A1.	
XX		
PD	08-OCT-1999.	
XX		
PF	07-APR-1998; 98FR-00004323.	
XX		
XX	07-APR-1998; 98FR-00004323.	
PR		
XX		
PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.	
PA	(CNRS) CNRS CENT NAT RECH SCI.	
PA	(INSP) INST PASTEUR LILLE.	
XX		
PI	Le Gal PA, Guillet JG, Gahery SH, Gras MR, Melnyk O, Tartar A;	
XX	WPI, 1999-583113/50.	
DR		
XX		
PT	New lipopeptide containing lipid regions and two epitopes, all separated	
PT	by peptide spacers that impart hydrophilicity, useful in vaccines.	
XX		
PS	Example 1; Page 9; 35pp; French.	
XX		
CC	The invention relates to the generation of a lipopeptide comprising at	
CC	least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)	
CC	epitope and at least one lipid residue with (i) the epitopes and lipid	
CC	portion and (ii) the epitopes, being separated independently by peptide	

CC spacers. These spacers comprise sequences of amino acids which carry an
CC overall electrical charge in neutral media to ensure that the lipopeptide
CC is hydrophilic. This peptide represents an example of them lipopeptide of
CC the invention. It contains 2 lipid residues attached at the N-terminal
CC end of the molecule and 2 epitopes separated by the spacer residues Ser-
CC Ser and Ala-Ala-Ala. Peptides AAV5301-Y53549 represent peptide epitopes
CC used in the generation of the lipopeptides. These are used in therapeutic
CC or prophylactic compositions and vaccines to induce specific immune
CC responses against human immunodeficiency, hepatitis B or papilloma
CC viruses; p53 of melanoma or the malaria parasite

SQ Sequence 28 AA;

Query Match 53.9%; Score 97; DB 2; Length 28;
Best Local Similarity 100.0%; Pred.No.3.9e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GROYIKANSKFGITERGR 19
Db 1 GROYIKANSKFGITERGR 19

RESULT 2
AAW71642
ID AAW71642 standard; protein; 280 AA.

AC AAW71642;

XX 17-OCT-2003 (revised)

DT 18-JAN-1999 (first entry)

XX HIV POL/NEF epitopes.

DE Vector; gene therapy; vaccine; ALVAC; translation factor; K3L; E3L;
KM vCP1433; HIV; pol; nef; epitope.

XX Human immunodeficiency virus 1.

XX MO9840500-A1.

XX 17-SEP-1998.

PD 25-FEB-1998; 98WO-US003710.

XX 12-MAR-1997; 97US-00815809.

XX (VIRO-) VIROGENETICS CORP.

PA (UYAR-) UNIV ARIZONA STATE.

XX Tartaglia J, Jacobs BL, Goebel SJ, Cox WI, Gettig RR, Pincus SE;
PI Paoletti E;

XX WPI; 1998-520819/44.

DR N-PSDB; AAV58243.

XX Enhancing expression of nucleic acids in cells - by using modified
PT vectors which comprise the nucleic acid and also nucleic acid encoding a
PT translation factor.

XX Example 1; Fig 4A-C; 90pp; English.

CC This polypeptide comprises the POL and NEF epitopes of HIV-1. It is
CC encoded by vCP1433 (see AAV58243). In vCP1433, an HIV pol/nef 'string of
CC beads' cassette is placed under control of the vaccinia H6 promoter.
CC pMPC6H6K3E3, containing a vaccinia H6/K3L expression cassette and
CC vaccinia E3L gene with endogenous promoter flanked by the ALVAC C6
CC insertion site sequences, was used in recombination with vCP1433 to
CC obtain vCP1452 (see AAV58244-45). K3L and E3L are vaccinia virus
CC translation factors. Novel vectors of the invention, such as ALVAC
CC vectors, include K3L and/or E3L and are used for enhancing expression of
CC gene products that they encode. The translation factors can effect
CC inhibition of eIF-2alpha phosphorylation or inhibition of protein kinase
CC PKR phosphorylation or otherwise sequester double stranded (ds) RNA,

CC increasing the effective concentration of ds RNA. The up-regulation of
CC foreign gene expression can have a profound effect on the induction of a
CC therapeutic or immunological response in a host administered or
CC inoculated with recombinants derived from these new vectors. (Updated on
CC 17-OCT-2003 to standardise OS field)

SQ Sequence 280 AA;

Query Match 49.4%; Score 89; DB 2; Length 280;
Best Local Similarity 77.3%; Pred.No.7.9e-05;
Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 13 GITERGRFPVTPQVPLRPMTYK 34
Db 145 GLTTMGFPVTPQVPLRPMTYK 166

RESULT 3
AAW71363
ID AAW71363 standard; protein; 280 AA.

AC AAW71363;

XX 17-OCT-2003 (revised)

DT 15-FEB-1999 (first entry)

XX HIV POL/NEF epitopes.

DE Vector; gene therapy; vaccine; ALVAC; K3L; E3L; translation factor;
KM vCP1433; HIV; POL; NEF; epitope.

XX Human immunodeficiency virus 1.

XX MO9840501-A1.

XX 17-SEP-1998.

PD 13-FEB-1998; 98WO-US002669.

XX 12-MAR-1997; 97US-00815155.

XX (VIRO-) VIROGENETICS CORP.

PA Tartaglia J, Cox WI, Gettig RR, Martinez H, Paoletti E;

XX WPI; 1998-520820/44.

DR N-PSDB; AAV60251.

XX Enhancing expression of nucleic acids in cells - by using modified
PT vectors which comprise nucleic acid and also nucleic acid encoding
PT transcription factor and optionally translation factor.

XX Example 2; Fig 6; 102pp; English.

CC This polypeptide comprises the POL and NEF epitopes of HIV-1. It is
CC encoded by vCP1433 (see AAV60251). In vCP1433, an HIV pol/nef 'string of
CC beads' cassette is placed under control of the vaccinia H6 promoter.
CC pMPC6H6K3E3, containing a vaccinia H6/K3L expression cassette and
CC vaccinia E3L gene with endogenous promoter flanked by the ALVAC C6
CC insertion site sequences, was used in recombination with vCP1433 to
CC obtain vCP1452 (see AAV60252-53). K3L and E3L are vaccinia virus
CC translation factors. New vectors are provided for enhanced expression of
CC at least 1 first nucleic acid molecule (NAM) in a cell having a
CC particular phenotype. The vector (e.g. NYAC or ALVAC) is modified to
CC comprise the first NAM and at least 1 second NAM encoding a transcription
CC factor (TF), or a TF and a translation factor such as K3L and E3L, where
CC there is co-temporal expression of the first and second NAMS with respect
CC to the phenotype of the cell, and where expression of the second NAM
CC enhances expression of the first NAM by enhancing transcription or
CC transcription and translation. Also claimed is a method for increasing
CC expression of at least 1 first NAM by such a vector. The vectors can be
CC used for increasing expression of e.g. an epitope of interest, a

CC induce an immune response. The lipoprotein induces a B cell and/or a T
 CC cell response. The lipoprotein is used to provide intranasal or
 CC sublingual immunization
 XX
 SQ Sequence 25 AA;

Query Match 48.6%; Score 87.5; DB 4; Length 25;
 Best Local Similarity 80.0%; Pred. No. 9.1e-06;
 Matches 20; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 GROYI-KANSKFIGITERGRFPVTP 24
 |||||
 1 GROYIKKANSKFIGITERGRILKEP 25

Db

RESULT 6
 AAY53551
 ID AAY53551 standard; protein; 29 AA.
 XX
 AC AAY53551;
 XX
 DT 18-JAN-2000 (first entry)
 XX
 DE Lipopeptide #2.
 XX
 KW Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
 KW electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
 KW human immunodeficiency virus; hepatitis B virus; papilloma virus;
 KW melanoma; malaria; parasite.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX

Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "contains palmitoyl residue attached at the N-terminus and on the epsilon carbon of the side chain"

FR2776926-A1.
 08-OCT-1999.
 07-APR-1998; 98FR-00004323.
 07-APR-1998; 98FR-00004323.
 (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 (CNRS) CNRS CENT NAT RECH SCI.
 (INSP) INST PASTEUR LILLE.
 Le Gal FA, Guillet JG, Gabery SH, Gras MH, Melnyk O, Tartar A;
 WPI; 1999-583113/50.
 New lipopeptide containing lipid regions and two epitopes, all separated
 by peptide spacers that impart hydrophilicity, useful in vaccines.
 Example 1; Page 9; 35pp; French.

The invention relates to the generation of a lipopeptide comprising at
 least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)
 epitope and at least one lipid residue with (i) the epitopes and lipid
 portion and (ii) the epitopes, being separated independently by peptide
 spacers. These spacers comprise sequences of amino acids which carry an
 overall electrical charge in neutral media to ensure that the lipopeptide
 is hydrophilic. This peptide represents an example of them lipopeptide of
 the invention. It contains 2 lipid residues attached at the N-terminal
 end of the molecule and 2 epitopes separated by the spacer residues Ser-
 Ser and Ala-Ala-Ala. Peptides AAY53301-Y53549 represent peptide epitopes
 used in the generation of the lipopeptides. These are used in therapeutic
 or prophylactic compositions and vaccines to induce specific immune
 responses against human immunodeficiency, hepatitis B or papilloma
 viruses; p53 of melanoma or the malaria parasite

XX
 SQ Sequence 29 AA;

Query Match 47.8%; Score 86; DB 2; Length 29;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 OYIKANSKFIGITERGR 19
 |||||
 4 OYIKANSKFIGITERGR 20

Db

RESULT 7
 AAP61515
 ID AAP61515 standard; protein; 206 AA.
 XX
 AC AAP61515;
 XX
 DT 27-AUG-2003 (revised)
 DT 08-JUN-1991 (first entry)
 XX
 DE Sequence of E' protein.
 XX
 KW HIV; IAV; AIDS; diagnosis; vaccine.
 OS Human T-cell lymphotropic virus 3B.
 XX
 PN BP187041-A.
 XX
 PD 09-JUL-1986.
 XX
 PF 23-DEC-1985; 85EP-00309454.
 XX
 PR 24-DEC-1984; 84US-00685272.
 PR 04-DEC-1985; 85US-00805069.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Capon DJ, Lasky LA;
 XX
 DR WPI; 1986-177602/28.
 DR N-PSDB; AAN60288.
 XX
 PT Acquired immune deficiency syndrome polypeptide(s) - obcd. by molecular
 PT cloning etc. and used for diagnosis and in vaccines against virus
 PT disease.
 XX
 PS Example; Fig 2; 125pp; English.
 XX
 CC A comparison of AAN60287 with the cDNA of the HTLV-III genome revealed
 CC one particular clone, designated p7.11 which contained a DNA sequence
 CC encoding this peptide (AAP60308) sequence. This approx. 2.2 kilobase
 CC covers the precursor gag region and encodes, 5' to 3', p-12, p-15, p-24 a
 CC second p-15 protein, and approx. 300 extra base pairs 3' to the gag
 CC region (see AAN60288). [Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 206 AA;

Query Match 47.2%; Score 85; DB 1; Length 206;
 Best Local Similarity 84.2%; Pred. No. 0.00022;
 Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMTYK 34
 |||||
 64 EKVGFPVTPQVPLRPMTYK 82

Db

RESULT 8
 AAM89326
 ID AAM89326 standard; protein; 206 AA.
 XX
 AC AAM89326;
 XX

DT 17-OCT-2003 (revised)
 DT 01-JUN-1999 (first entry)
 XX
 DE HIV-1 nef protein sequence.
 XX
 KW Antigenic composition; primate; lentivirus; nef gene; vaccine; infection;
 KW AIDS; HIV-1; nef protein.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN US5851813-A.
 XX
 PD 22-DEC-1998.
 XX
 PF 27-JAN-1994; 94US-00186583.
 XX
 PR 12-JUL-1990; 90US-00551945.
 PR 09-JUL-1991; 91US-00727494.
 XX
 PA (HARD) HARVARD COLLEGE.
 XX
 PI Destrosiers RC;
 XX
 DR MPI; 1999-080408/07.
 DR N-PSDB; AAV81866.
 XX
 PT Lentivirus antigenic compositions - containing lentivirus with nef gene
 PT deletion.
 XX
 PS Disclosure; Fig 2A-R; 93pp; English.
 XX
 CC The invention relates to an antigenic composition comprising an isolated
 CC primate lentivirus whose genome contains an engineered non-reversible
 CC null mutation in the nef gene, or an infectious DNA clone in a carrier.
 CC The antigenic composition is used in vaccines against infection by the
 CC lentivirus, e.g. AIDS. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 206 AA;

Query Match 47.2%; Score 85; DB 2; Length 206;
 Best Local Similarity 84.2%; Pred. No. 0.00022;
 Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGPPVTPQVPLRPMYK 34
 : |||||
 Db 64 EKVGPVTPQVPLRPMYK 82

RESULT 9
 ABU63323
 ID ABU63323 standard; protein; 206 AA.
 XX
 AC ABU63323;
 XX
 DT 24-SEP-2003 (first entry)
 XX
 DE Human lymphotropic virus type III E' polypeptide.
 XX
 KW Vaccine; E'; AIDS; AIDS-associated retrovirus; blood testing;
 KW retrovirus dissemination inhibition; blood product testing.
 XX
 OS Human lymphotropic virus type III.
 XX
 PN US6534285-B1.
 XX
 PD 18-MAR-2003.
 XX
 PF 12-APR-2000; 2000US-00547692.
 XX
 PR 24-DEC-1984; 84US-00685272.
 PR 04-DEC-1985; 85US-00805069.
 PR 08-MAY-1986; 86US-00861016.
 PR 02-AUG-1988; 88US-00227568.

PR 19-NOV-1992; 92US-00979391.
 PR 29-SEP-1993; 93US-00129009.
 PR 29-JUL-1994; 94US-00282857.
 PR 29-JUL-1998; 98US-00124596.
 XX
 PA (GENTECH) GENENTECH INC.
 XX
 PI Berman PW, Capon DJ, Lasky LA;
 XX
 DR MPI; 2003-531076/50.
 DR N-PSDB; ACD27911.
 XX
 PT Composition useful as vaccine, has AIDS-associated retrovirus polypeptide
 PT having antigenic determinant that specifically binds to complementary
 PT antibody, and is free of other AIDS-associated polypeptides.
 XX
 PS Example 1; Fig 2; 49pp; English.
 XX
 CC The invention relates to a composition which comprises a predetermined
 CC polypeptide sequence of an AIDS-associated retrovirus (AR), or its
 CC fragment, where AR or its fragment comprises at least one antigenic
 CC determinant that specifically binds to complementary antibody, and is
 CC essentially free of other naturally occurring AIDS-associated polypeptide
 CC sequences or human proteins from cells for which AR is naturally
 CC infective. The composition is useful as diagnostic products and vaccines
 CC in the detection of and vaccination against viral aetiological agents of
 CC AIDS. The composition is also useful for inhibiting infection by AIDS-
 CC associated retrovirus and dissemination of retrovirus in infected
 CC individuals. The AR is useful in an assay system to identify compounds
 CC which inhibit AIDS-associated reverse transcriptase. The AR is also
 CC useful as diagnostic agents to detect AIDS in individuals donated blood
 CC and blood products and as immunogens in the production of neutralising
 CC antibodies which confer resistance to infection by AIDS-associated
 CC retrovirus. The present sequence represents the amino acid sequence of
 CC the human lymphotropic virus type III E' polypeptide
 XX
 SQ Sequence 206 AA;

Query Match 47.2%; Score 85; DB 6; Length 206;
 Best Local Similarity 84.2%; Pred. No. 0.00022;
 Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGPPVTPQVPLRPMYK 34
 : |||||
 Db 64 EKVGPVTPQVPLRPMYK 82

RESULT 10
 ABR44739
 ID ABR44739 standard; protein; 505 AA.
 XX
 AC ABR44739;
 XX
 DT 28-JUL-2003 (first entry)
 XX
 DE Plasmid p17/24opt/rNef1 insert protein sequence SEQ ID NO:55.
 XX
 KW Human immunodeficiency virus; HIV-1; vaccine; gene therapy;
 KW acquired immune deficiency syndrome; AIDS; HIV infection.
 XX
 OS Human immunodeficiency virus 1.
 OS Synthetic.
 XX
 PN WO2003025003-A2.
 XX
 PD 27-MAR-2003.
 XX
 PF 18-SEP-2002; 2002WO-EP010592.
 XX
 PR 20-SEP-2001; 2001WO-GB004207.
 PR 11-DEC-2001; 2001GB-00029604.
 PR 19-MAR-2002; 2002GB-00006462.
 XX

PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Beaton A, Ertl PF, Gough GW, Lear A, Tite JP, Van Wely CA;
 XX
 XX MPI: 2003-363125/34.
 DR N-PSDB; ACC69912.
 XX
 PT New nucleotide sequences encoding an HIV-1 gag protein and a second HIV
 PT antigen, or their fragments, useful in medicine, particularly as vaccines
 PT for treating acquired immune deficiency syndrome (AIDS) or HIV
 PT infections.
 PS Example 3; Fig 4; 116pp; English.
 XX
 XX The present invention describes a nucleotide sequence encoding an HIV-1
 CC gag protein or its fragment containing a gag epitope, and a second HIV
 CC antigen or its fragment encoding an epitope of the second HIV antigen,
 CC operably linked to a heterologous promoter. Also described: (1) a vector
 CC comprising the nucleotide sequence; (2) a pharmaceutical composition
 CC comprising the nucleotide sequence or the vector, and a pharmaceutical
 CC excipient, diluent, carrier or adjuvant; (3) an intradermal delivery
 CC device comprising the pharmaceutical composition; and (4) a process for
 CC the production of the nucleotide. HIV epitopes and antigens have virucide
 CC activity, and can be used in vaccines and gene therapy. The nucleotide
 CC sequence, vector or pharmaceutical compositions can be used in medicine,
 CC or for treating a patient suffering from or susceptible to a disease. The
 CC nucleotide sequence is also useful for manufacturing a medicament for the
 CC treatment of disease. The nucleic acid is particularly useful as DNA
 CC vaccines for treating e.g. acquired immune deficiency syndrome (AIDS) or
 CC HIV infections. The present sequence represents an HIV plasmid amino acid
 CC sequence, which is used in an example from the present invention
 CC
 SQ Sequence 505 AA;
 XX
 QY
 Query Match 46.7%; Score 84; DB 6; Length 505;
 Best Local Similarity 62.1%; Pred. No. 0.00084;
 Matches 18; Conservative 1; Mismatches 4; Indels 6; Gaps 1;
 DB 359 KARVIMVG-----FVTPQVPLRPMTYK 381
 || : |||||
 6 KANSKFIGITGRGPPVTPQVPLRPMTYK 34
 || : |||||
 359 KARVIMVG-----FVTPQVPLRPMTYK 381
 RESULT 11
 ID ABR44747 standard; protein; 505 AA.
 XX
 AC ABR44747;
 XX
 DT 28-JUL-2003 (first entry)
 XX
 DE Plasmid p73i-GN2 insert protein sequence SEQ ID NO:71.
 XX
 DE Human immunodeficiency virus; HIV-1; vaccine; virucide; gene therapy;
 KW acquired immune deficiency syndrome; AIDS; HIV infection.
 XX
 XX Human immunodeficiency virus 1.
 OS Synthetic.
 OS
 OS WO2003025003-A2.
 XX
 PN 27-MAR-2003.
 XX
 PD 18-SEP-2002; 2002WO-EP010592.
 PF
 XX 20-SEP-2001; 2001WO-GB004207.
 XX PR 11-DEC-2001; 2001GB-00029604.
 PR 19-MAR-2002; 2002GB-00006462.
 XX
 PA (GLAX) GLAXO GROUP LTD.
 PI Beaton A, Ertl PF, Gough GW, Lear A, Tite JP, Van Wely CA;
 XX

DR MPI: 2003-363125/34.
 DR N-PSDB; ACC69912.
 XX
 PT New nucleotide sequences encoding an HIV-1 gag protein and a second HIV
 PT antigen, or their fragments, useful in medicine, particularly as vaccines
 PT for treating acquired immune deficiency syndrome (AIDS) or HIV
 PT infections.
 PS Example 11; Fig 12; 116pp; English.
 XX
 XX The present invention describes a nucleotide sequence encoding an HIV-1
 CC gag protein or its fragment containing a gag epitope, and a second HIV
 CC antigen or its fragment encoding an epitope of the second HIV antigen,
 CC operably linked to a heterologous promoter. Also described: (1) a vector
 CC comprising the nucleotide sequence; (2) a pharmaceutical composition
 CC comprising the nucleotide sequence or the vector, and a pharmaceutical
 CC excipient, diluent, carrier or adjuvant; (3) an intradermal delivery
 CC device comprising the pharmaceutical composition; and (4) a process for
 CC the production of the nucleotide. HIV epitopes and antigens have virucide
 CC activity, and can be used in vaccines and gene therapy. The nucleotide
 CC sequence, vector or pharmaceutical compositions can be used in medicine,
 CC or for treating a patient suffering from or susceptible to a disease. The
 CC nucleotide sequence is also useful for manufacturing a medicament for the
 CC treatment of disease. The nucleic acid is particularly useful as DNA
 CC vaccines for treating e.g. acquired immune deficiency syndrome (AIDS) or
 CC HIV infections. The present sequence represents an HIV plasmid amino acid
 CC sequence, which is used in an example from the present invention
 CC
 SQ Sequence 505 AA;
 XX
 QY
 Query Match 46.7%; Score 84; DB 6; Length 505;
 Best Local Similarity 62.1%; Pred. No. 0.00084;
 Matches 18; Conservative 1; Mismatches 4; Indels 6; Gaps 1;
 DB 359 KARVIMVG-----FVTPQVPLRPMTYK 381
 || : |||||
 6 KANSKFIGITGRGPPVTPQVPLRPMTYK 34
 || : |||||
 359 KARVIMVG-----FVTPQVPLRPMTYK 381
 RESULT 12
 ID ABR44751 standard; protein; 1067 AA.
 XX
 AC ABR44751;
 XX
 DT 28-JUL-2003 (first entry)
 XX
 DE Plasmid Trgn #6 insert protein sequence SEQ ID NO:80.
 XX
 DE Human immunodeficiency virus; HIV-1; vaccine; virucide; gene therapy;
 KW acquired immune deficiency syndrome; AIDS; HIV infection.
 XX
 XX Human immunodeficiency virus 1.
 OS Synthetic.
 OS
 OS Key Location/Qualifiers
 FT Misc-difference 11 /note="encoded by CCG"
 FT FT
 FT Misc-difference 231 /note="encoded by AAG"
 FT FT
 FT
 XX
 PN 27-MAR-2003.
 XX
 PD 18-SEP-2002; 2002WO-EP010592.
 PF
 XX 20-SEP-2001; 2001WO-GB004207.
 XX PR 11-DEC-2001; 2001GB-00029604.
 PR 19-MAR-2002; 2002GB-00006462.
 XX
 PA (GLAX) GLAXO GROUP LTD.
 PI Beaton A, Ertl PF, Gough GW, Lear A, Tite JP, Van Wely CA;
 XX

PI	Beaton A, Ertl PF, Gough GW, Lear A, Tite JP, Van Wely CA;
XX	
DR	WPI, 2003-363125/34.
DR	N-PSDB; ACC69917.
XX	
XX	
PT	New nucleotide sequences encoding an HIV-1 gag protein and a second HIV
PT	antigen, or their fragments, useful in medicine, particularly as vaccines
PT	for treating acquired immune deficiency syndrome (AIDS) or HIV
PT	infections.
XX	
PS	
PS	Example 16; Fig 17; 116pp; English.
XX	
CC	
CC	The present invention describes a nucleotide sequence encoding an HIV-1
CC	gag protein or its fragment containing a gag epitope, and a second HIV
CC	antigen or its fragment encoding an epitope of the second HIV antigen,
CC	operably linked to a heterologous promoter. Also described: (1) a vector
CC	comprising the nucleotide sequence; (2) a pharmaceutical composition
CC	comprising the nucleotide sequence or the vector, and a pharmaceutical
CC	excipient, diluent, carrier or adjuvant; (3) an intradermal delivery
CC	device comprising the pharmaceutical composition; and (4) a process for
CC	the production of the nucleotide. HIV epitopes and antigens have virulide
CC	activity, and can be used in vaccines and gene therapy. The nucleotide
CC	sequence, vector or pharmaceutical compositions can be used in medicine,
CC	or for treating a patient suffering from or susceptible to a disease. The
CC	nucleotide sequence is also useful for manufacturing a medicament for the
CC	treatment of disease. The nucleic acid is particularly useful as DNA
CC	vaccines for treating e.g. acquired immune deficiency syndrome (AIDS) or
CC	HIV infections. The present sequence represents an HIV plasmid amino acid
CC	sequence, which is used in an example from the present invention
XX	
SQ	Sequence 1067 AA:
Query Match	46.7%; Score 84; DB 6; Length 1067;
Best Local Similarity	62.1%; Fred. No. 0.0019;
Matches 18; Conservative 1; Mismatches 4; Indels 6; Gaps 1;	
QY	6 KANSKFIGITGRGFPVTPQVPLRPMYK 34
	:
DB	921 KARVLMVG-----FPVTPQVPLRPMYK 943
RESULT 13	
ABR44753	
ID	ABR44753 standard; protein; 1067 AA.
XX	
AC	ABR44753;
XX	
DT	28-JUN-2003 (first entry)
XX	
DE	
XX	Plasmid Tngr(F1) protein sequence SEQ ID NO:84.
KM	Human immunodeficiency virus; HIV-1; vaccine; virucide; gene therapy;
KM	acquired immune deficiency syndrome; AIDS; HIV infection.
XX	
OS	Human immunodeficiency virus 1.
XX	
XX	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 516
FT	/note= "encoded by CCG"
FT	Misc-difference 736
FT	/note= "encoded by AAG"
XX	
PN	WO2003025003-A2.
XX	
PD	
XX	27-MAR-2003.
XX	
XX	18-SEP-2002; 2002WO-EP010592.
XX	
XX	20-SEP-2001; 2001WO-GB004207.
PR	11-DEC-2001; 2001GB-00029604.
PR	19-MAR-2002; 2002GB-00006462.
XX	

```

PA      (GLAXO ) GLAXO GROUP LTD.
XX
PI      Beaton A, Ertl PF, Gough GW, Lear A, Tite JP, Van Wely CA;
XX
DR      WPI, 2003-363125/34.
XX
DR      N-PESDB; ACC69919.
XX
PT      New nucleotide sequences encoding an HIV-1 gag protein and a second HIV
PT      antigen, or their fragments, useful in medicine, particularly as vaccines
PT      for treating acquired immune deficiency syndrome (AIDS) or HIV
PT      infections.
XX
PS      Example 18; Fig 19; 116pp; English.
XX
CC      The present invention describes a nucleotide sequence encoding an HIV-1
CC      gag protein or its fragment containing a gag epitope, and a second HIV
CC      antigen or its fragment encoding an epitope of the second HIV antigen,
CC      operably linked to a heterologous promoter. Also described: (1) a vector
CC      comprising the nucleotide sequence; (2) a pharmaceutical composition
CC      comprising the nucleotide sequence or the vector, and a pharmaceutical
CC      excipient, diluent, carrier or adjuvant; (3) an intradermal delivery
CC      device comprising the pharmaceutical composition; and (4) a process for
CC      the production of the nucleotide. HIV epitopes and antigens have virucide
CC      activity, and can be used in vaccines and gene therapy. The nucleotide
CC      sequence, vector or pharmaceutical compositions can be used in medicine,
CC      or for treating a patient suffering from or susceptible to a disease. The
CC      nucleotide sequence is also useful for manufacturing a medicament for the
CC      treatment of disease. The nucleic acid is particularly useful as DNA
CC      vaccines for treating e.g. acquired immune deficiency syndrome (AIDS) or
CC      HIV infections. The present sequence represents an HIV plasmid amino acid
CC      sequence, which is used in an example from the present invention
XX
SQ      Sequence 1067 AA.
XX
Query Match      46.7%; Score 84; DB 6; Length 1067;
Best Local Similarity      62.1%; Pred. No. 0.0019;
Matches      18; Conservative      1; Mismatches      4; Indels      6; Gaps      1.
OY      6 KANSKEFIGITERGRFPVTPPOVLRPMPTK 34
      || -|| ||||| ||||| |||||
DB      359 KARVLWVG-----FPVTPPOVLRPMPTK 381

RESULT 14
AAB31115
ID      AAB31115 standard; peptide; 30 AA.
AC      AAB31115;
XX
DT      02-APR-2001 (first entry)
XX
DE      A polypeptidic fragment of HIV Nef, comprising residues 68-97.
XX
KW      Polypeptidic peptide; p53; cancer; human leukocyte antigen; HLA;
KW      immune response; cytolysis; cytotoxic T cell; CTL; cytokine secretion;
KW      interleukin-2; IL-2; IL-4; gamma-interferon; p53-related cancer; NEF;
KW      HIV.
XX
OS      Human immunodeficiency virus.
XX
PN      FR2794368-A1.
XX
PD      08-DEC-2000.
XX
PF      07-OCT-1999; 99FR-00012512.
XX
PR      03-JUN-1999; 99FR-00007012.
XX
PA      (BIOV-) BIOVECTOR THERAPEUTICS SA.
PA      (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX      Chopin J, Bourgault VI, Guillet JG, Connan F, Ferries E;
XX

```

DR WPI; 2001-064173/08.
 XX
 PT New polypeptidic fragments from the p53 protein, useful for treatment or
 PT prevention of cancer, e.g. of breast or colon.
 XX
 PS Disclosure; Page 6; 26pp; French.
 XX
 CC The present sequence represents a polypeptidic fragment of HIV Nef
 CC protein. The specification describes polypeptidic fragments of human p53,
 CC which is overexpressed in many types of cancers. The p53 polypeptidic
 CC fragments bind stably to human leukocyte antigen (HLA) type molecules.
 CC The p53 peptides induce a specific immune response. They induce
 CC cytotoxic T cells (CTL) of cells that express the p53
 CC peptides associated with appropriate HLA molecules and induce secretion
 CC of cytokines (particularly interleukin (IL)-2 and IL-4, and gamma-
 CC interferon) by these CTL. The p53 peptides, derivatives, nucleic acids
 CC encoding them and specific antibodies are used, in compositions or
 CC vaccines, to treat or prevent diseases p53-related cancers, particularly
 CC of breast, colon, lung or bladder
 XX
 SQ Sequence 30 AA;
 Query Match 46.1%; Score 83; DB 4; Length 30;
 Best Local Similarity 100.0%; Pred. No. 5.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 FPVTPQVPLRPMTYK 34
 |||||
 1 FPVTPQVPLRPMTYK 15
 DB
 RESULT 15
 AAB30938
 ID AAB30938 standard; peptide; 30 AA.
 XX
 AC AAB30938;
 XX
 DT 02-APR-2001 (first entry)
 XX
 DE Peptide comprising amino acids 68-97 of HIV Nef protein.
 XX
 KW Polypeptidic peptide; E6 protein; HPV; CD4 epitope;
 KW T helper cell; human leukocyte antigen; HLA; immune response; cytotoxic
 KW cytotoxic T cell; CTL; cytokine secretion; interleukin-2; IL-2; IL-4;
 KW gamma-interferon; HPV infection; cervical neoplasia; invasive cancer;
 KW vulvar intraepithelial neoplasia; Nef.
 XX
 OS Human immunodeficiency virus.
 XX
 PN FR2794371-A1.
 PD 08-DEC-2000.
 XX
 PF 07-OCT-1999; 99FR-00012511.
 XX
 PR 03-JUN-1999; 99FR-00007012.
 XX
 PA (BIOV-) BIOVECTOR THERAPEUTICS SA.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX
 PI Choppin J, Bourgault VI, Guillet JG, Connan F, Ferries E;
 XX
 WPI; 2001-064175/08.
 XX
 PT New polypeptidic fragments from human papilloma virus E6 and E7 proteins,
 PT useful for treatment or prevention of e.g. cervical neoplasia and cancer.
 XX
 PS Disclosure; Page 6; 27pp; French.
 XX
 CC The present sequence is derived from the Nef protein of Human
 CC immunodeficiency virus (HIV). The peptide is an polypeptidic fragment of
 CC Nef. The specification describes polypeptidic fragments from the E6 and
 CC E7 proteins of human papilloma virus (HPV). The HPV peptides include CD4

CC epitopes recognised by T helper cells. They bind stably to human
 CC leukocyte antigen (HLA) type molecules. The HPV peptides induce a
 CC specific immune response, particularly cytotoxic T, caused by cytotoxic T
 CC cells (CTL). They also induce secretion of cytokines (particularly
 CC interleukin-2 (IL-2) and IL-4, and gamma-interferon) by CTL. The HPV
 CC peptides, their derivatives, nucleic acids encoding them and specific
 CC antibodies are used, in compositions or vaccines, to treat or prevent
 CC diseases associated with HPV infection, e.g. cervical or vulvar
 CC intraepithelial neoplasia and invasive cancer of the cervix uteri. The
 CC antibodies are also useful for in vitro diagnosis of these diseases
 XX
 SQ Sequence 30 AA;
 Query Match 46.1%; Score 83; DB 4; Length 30;
 Best Local Similarity 100.0%; Pred. No. 5.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 FPVTPQVPLRPMTYK 34
 |||||
 1 FPVTPQVPLRPMTYK 15
 DB

Search completed: July 20, 2004, 06:30:29
 Job time : 51.7213 secs